UNIVERSITY OF LONDON THESIS

Degree PhD Year 2008 Name of Author KING, Christopher William

A New Approach

to

Stitching Optical Metrology Data.

by

Christopher W. King

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

OF THE UNIVERSITY OF LONDON

Department of Physics and Astronomy

UNIVERSITY COLLEGE LONDON

University of London

September 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

To whom it may concern,

I, Christopher W. King, hereby declare that all work presented in this thesis, unless stated otherwise, is my own.

Abstract

The next generation of optical instruments, including telescopes and imaging apparatus, will generate an increased requirement for larger and more complex optical forms. A major limiting factor for the production of such optical components is the metrology: how do we measure such parts and with repect to what reference datum? This metrology can be thought of as part of a complete cycle in the production of optical components and it is currently the most challenging aspect of production.

This thesis investigates a new and complete approach to stitching optical metrology data to extend the effective aperture or, in future, the dynamic range of optical metrology instruments. A practical approach is used to build up a complete process for stitching on plano and spherical parts. The work forms a basis upon which a stitching system for aspheres might be developed in the future, which is inherently more complicated.

Beginning with a historical perspective and a review of optical polishing and metrology, the work presented relates the commercially available metrology instruments to the stitching process developed. The stitching is then performed by a numerical optimization routine that seeks to join together overlapping sub-aperture measurements by consideration of the aberrations introduced by the measurement

scenario, and by the overlap areas between measurements.

The stitching is part of a larger project, the PPARC Optical Manipulation and Metrology project, and was to benefit from new wavefront sensing technology developed by a project partner, and to be used for the sub-aperture measurement. Difficult mathematical problems meant that such a wavefront sensor was not available for this work and a work-around was therefore developed using commercial instruments. The techniques developed can be adapted to work on commercial machine platforms, and in partuicular, the OMAM NPL/UCL swing-arm profilometer described in chapter 5, or the computer controlled polishing machines as manufactured by the project's industrial sponsors, Zeeko ltd.

The novel process involving physical surface fiducialization is then developed for a standpoint of instrument particularities followed by the development of a numerical optimization to derive the *best* stitching. The software, developed in Matlab, is then tested on simulated and real data and shows much promise. The stitching accuracy is found to be dictated by the quality of the acquired data and the precision with which fiducials in a given measurement can be related to each other.

Finally a summary is given for a future direction of this work and how it might be improved upon.

Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

Writing this thesis has involved the help and support of many people, and I would like to express my sincere thanks to all those who have aided the process. In particular, to Dr David Walker for supervising the project, and also to many of the staff and students of Group A in physics and astronomy at UCL. I would like to thank Dr Peter Doel for his proofreading and discussions on metrology and wavefront sensing issues, and Dr David Brooks for his help, humour and advice with the practices of optical metrology.

Thanks are also due to my many friends, some of whom are colleagues at UCL, and some of whom I have not seen for a while.

The work was been made possible by grants from the EPSRC and from Zeeko Ltd, the industrial sponsor, for which I am extremely grateful.

Special gratitude goes to my parents who have supported me throughout my academic and professional career. I would not have been able to pursue my path without their help and support.

COMMERCIAL IN CONFIDENCE

CHAPTER 1

Introduction

"It is difficult to say what is impossible, for the dream of yesterday is the hope of today and the reality of tomorrow."

– Robert H. Goddard, physicist and rocket engineer (1882 - 1945)

1.1 Project Context

The present trend and need for large, ultra-precise optical surfaces is causing a revolution in the techniques being applied in the field of optical manufacture and metrology. Deterministic polishing is now possible thanks to several novel processes of recent development. Metrology, however, remains a significant source of problems to optical designers and manufacturers. Many of the modern polishing technologies summarized later in on section 1.5 are well suited to the production of non-spherical surfaces, but they are currently limited by a lack of accurate metrology. Problems occur when long optical path lengths, fast components, aspheric components, freeform components and large-aperture parts are encountered.

The problems of producing precise, spherical and non-spherical optical and me-

chanical components are encountered in many areas including research optics production, prosthetic joint production and industrial mould production. The requirements and tolerances on surface forms become ever tighter on ever more complex surfaces and the traditional optical tests in use for many years become obsolete and require augmentation or replacement if a sufficiently precise and accurate measurement is to be made on such forms.

This work investigates a novel approach to stitching optical metrology data for use on large aperture optical surfaces (>150 mm diameter), high numerical aperture spheres, and in future, on very slow off axis aspheres as required for future telescope projects. Stitching simply involves joining overlapping measurements of parts of a surface together to produce a map of the whole surface.

Previous work, summarized in section 1.11, has been reported in the past on stitching metrology data for plano objects and spheres using very accurate and expensive CNC (Computer Numerical Control) hardware platforms. This work avoids the requirement for such hardware and uses prior knowledge within overlapping surface patches to stitch data sets together. This prior knowledge is gained by placing removable opaque ink spots onto the part to be tested. These spots provide physical features in the metrology data to allow stitching to occur - a points based registration technique.

The requirement for stitching comes from the need to verify, and eventually replace, traditional surface metrology on large aperture and often aspheric parts. Taking the example of a large off-axis paraboloidal mirror of long focal length, we could devise a full-aperture interferometric test and obtain a null measurement by using a test tower, a return flat and some form of wavefront sensor. As the part has a long focal length, air turbulence, air stratification and vibration immediately cause problems. The long focal length also requires the construction and metrology of an expensive test tower to do the test in the traditional way using a return flat

and testing at the centre of curvature.

Stitching, in principle, may be used to perform a much more compact test of such an optical surface, eliminating the path length problems, with the added benefit of achieving a higher spatial resolution than a single full-aperture test – the same number of measurement points is used over a smaller portion of the surface. These points are more fully explored in the later chapters of this work.

The most basic gains brought by stitching are:

- Increased spatial resolution of measurement over the full-aperture of a part when compared to a full-aperture test.
- Parts whose apertures exceed the capacity of traditional full-aperture testing instruments can be tested using sub-aperture methods.
- Shorter optical path lengths can be used to eliminate problems caused by air turbulence and vibration.
- Parts with aspheric departures too large for full-aperture testing without null correctors can be tested using sub-aperture methods.

In this thesis a new stitching algorithm is developed to stitch optical metrology data taking into account the problems encountered when particular types of instruments are used. The thesis begins with a review of optical production and then moves on the the specific advantages of stitching before the method is developed and demonstrated.

1.2 A Brief History of Progress in Optics

1.2.1 Ancient Greece and Classical Antiquity

For hundreds of years man has attempted to produce optical devices for various uses and the earliest examples predate classical antiquity. In 1914 an unbroken mirror

was found in the workers quarter of the pyramid of Sesostris II in the Nile Valley dating to 1900 BC, evidence that the ancient Egyptians knew how to make mirrors. Years later Aristophanes, whilst lampooning the sophists and the intellectuals in 5th century Athens, mentioned a *burning glass* in his comic play '*The Clouds*' in approximately 424 BC. Indeed, our modern term *optics* has its origins in the ancient Greek for *appearance* or *look*, '*όπτικήν*'.

In the 3rd century BC, Euclid of Alexandria used his principles of geometry to write an early treatise on optics called '*Catoptrics*'. In it Euclid discussed mirrors, the effects of perspective and seeing things at a distance. However, he did not really discuss the nature of the visible rays of light; in common with other early writers, Euclid treated vision as a geometrical rather than a physical or psychological problem. His study began in a similar fashion to his work '*Elements*' on geometry with a set of self-evident axioms [2]:

1. Lines (or visual rays) can be drawn in a straight line to the object.
2. Those lines falling upon an object form a cone.
3. Those things upon which the lines fall are seen.
4. Those things seen under a larger angle appear larger.
5. Those things seen by a higher ray, appear higher.
6. Right and left rays appear right and left.
7. Things seen within several angles appear clearer.

Not all the axioms were correct, but they did help him understand certain aspects of vision that were widely misunderstood at the time.

Claudius Ptolemaeus (Ptolemy, c.90–168 AD) [3], the famous Greek mathematician, geographer, astrologer and astronomer also wrote a treatise on optics covering his theories of reflection, refraction and colour. Today, only an Arabic translation of the text exists that was used by followers of Ptolemy to study optics. Unfortunately, Euclid and Ptolemy made a common mistake in that they assumed that visual rays

proceeded from the eye to the object seen – a statement easily repudiated; it was to be some 700 years later until this was contradicted in another work. Of course, Ptolemy is most famous for his '*Amalgest*' [4] - great book - on mathematics and astronomy where he wrote extensively on the theories of the Greek and Babylonian world.

1.2.2 The Islamic world

The pre-eminence of the Greek and Roman scholars eventually gave way to that of the Islamic scholars, particularly after the fall of the Roman Empire. Al-Kindi (801-873) played an important role in mathematics and optics and also in more miscellaneous subjects such as the manufacture of swords [5]. He knew of the Aristotelean and Euclidean theories of vision and perception. He also knew which was better at explaining the process of vision from his own experiences. Aristotle's theory of vision could not account for how the object observed was perceived when viewed at an angle. For example, a square will appear as a line when viewed side-on, and Euclid's theory could fully account for this by the provision of a geometrical model assuming that light could travel only in straight lines. Euclid's model could also explain the action of mirrors and the lengths of shadows. Al-Kindi was a supporter and extender of the Euclidean theory of optics at the time. He would prove to be influential in the works of other scholars and he went on to write approximately 260 books including 32 on geometry and 12 on physics.

In 984 [7] another scholar attached to the court of Baghdad, Ibn Sahl, wrote a treatise on '*Burning Mirrors and Lenses*' in which he set out his understanding of catoptrics and dioptrics or the reflection and refraction of light. It is probable that Rashed [8] called Sahl 'A pioneer of anaclastics' i.e., a pioneer of aspherical optics, for his study of conic forms. Sahl's treatise '*On the Burning Instruments*' was concerned primarily with burning at a point using a nearby or distant luminous

Figure 1.1: Al-Kindi, an Arab scholar, wrote extensively on optics and mathematics during the 9^{th} century. He was an advocate of the Euclidean optics. [6]

source by reflection of refraction. According to Rashed, Sahl studied conical forms for reflecting and refracting media and is credited with the discovery of the law of refraction known as Snell's law, some 637 years before Willebrord Snellius.

Alhazen (965–1040) [9] took up the mantle after Sahl and wrote his *Kitab al-Manazir* (book on optics) between 1015 and 1021. In it Alhazen agreed with Aristotle that we could see because light originates from the objects we observe, and not that the rays emanate from our eyes as Euclid and Ptolemy had stated. Alhazen was thus able to establish a successful theory to explain the process of vision from the point of view that rays reached the eye from each point of the object under observation. Alhazen is thus credited with formulating the first alternative and comprehensive theory of vision to the earlier Greeks. He was also able to demonstrate that light and colour could pass through a single aperture in a straight line without intermingling using an early *camera obscura*.

1.2.3 Western Middle Ages

Europe entered a period often referred to as the Dark Ages after the collapse of the Roman Empire, where scientific progress was slow. Much of the science developed was concerned with logic, empiricism and the scientific method. A few notables studied optics including the English Bishop Robert Grosseteste and the Franciscan Roger Bacon whom many regard as the father of modern scientific study. Bacon [10] drew from the work of Ptolemy, Al-Kindi and Alhazen but progress still remained slow until the beginnings of the 16^{th} century.

1.2.4 The Scientific Revolution

The year 1543 saw the publication of '*De Revolutionibus Orbium Coelestium*' (On the Revolutions of the Celestial Spheres) [11] by Nicolaus Copernicus, and is widely regarded as marking the beginning of the scientific revolution. Shortly after Copernicus's publication an obscure Dutch spectacle maker named Hans Lippershey applied for a patent in 1608 on the telescope, which he called a kijker (*looker*) [12], beating Jacob Metius by a few weeks. The patent was not granted but the information was disseminated around Europe where it soon reached Italy.

Galileo Galilei (1564–1619) heard about the *looker* from Jacques de Bovedere of Paris [12] and undertook to construct his own version by grinding his own lenses and using a positive object lens and a negative eyelens – what we now call a Galilean telescope. It is now well known that Galileo was the first to point a telescope at the night sky and to record and disseminate his observations in a new and striking way in his *Siderus Nuncius* or Sidereal Messenger in 1610.

The refracting telescope reigned for many years to come, with several improvements along the way including the use of a positive eyelens (attributed to Kepler) amongst others. It was Newton, though, who developed the first successful reflecting telescope in 1668 after concluding that it was impossible to remove chromatic

Figure 1.2: Galileo Galilei by Justus Sustermans, 1639, Galileo was the first person to study the heavens by telescope, from [13].

aberration from refracting telescopes. He developed a procedure for grinding speculum alloy into spherical mirrors and used an interference technique now known as Newton's rings to assess the quality of the optics. His telescope had a much wider aperture than contemporary refractors and he was thus able to build a superior instrument that was first demonstrated to the Royal society in 1671 [14].

In 1637 [15] René Descartes published the now-familiar form of Snell's law which he deduced by using a model that treated light as a pressure transmitted by an elastic medium. Pierre de Fermat later re-derived the law using his own principle of *least-time* in 1657 [15]. From this point on our knowledge and understanding of optics increased rapidly, from diffraction by Grimaldi and others, to Maxwell's equations on the eve of the 20^{th} century. A summary of some key dates in the development of optics is shown in figure 1.3. The width and colour of the colour bars is proportional to the number of years between entries, with 0.1 mm equating to 1 year.

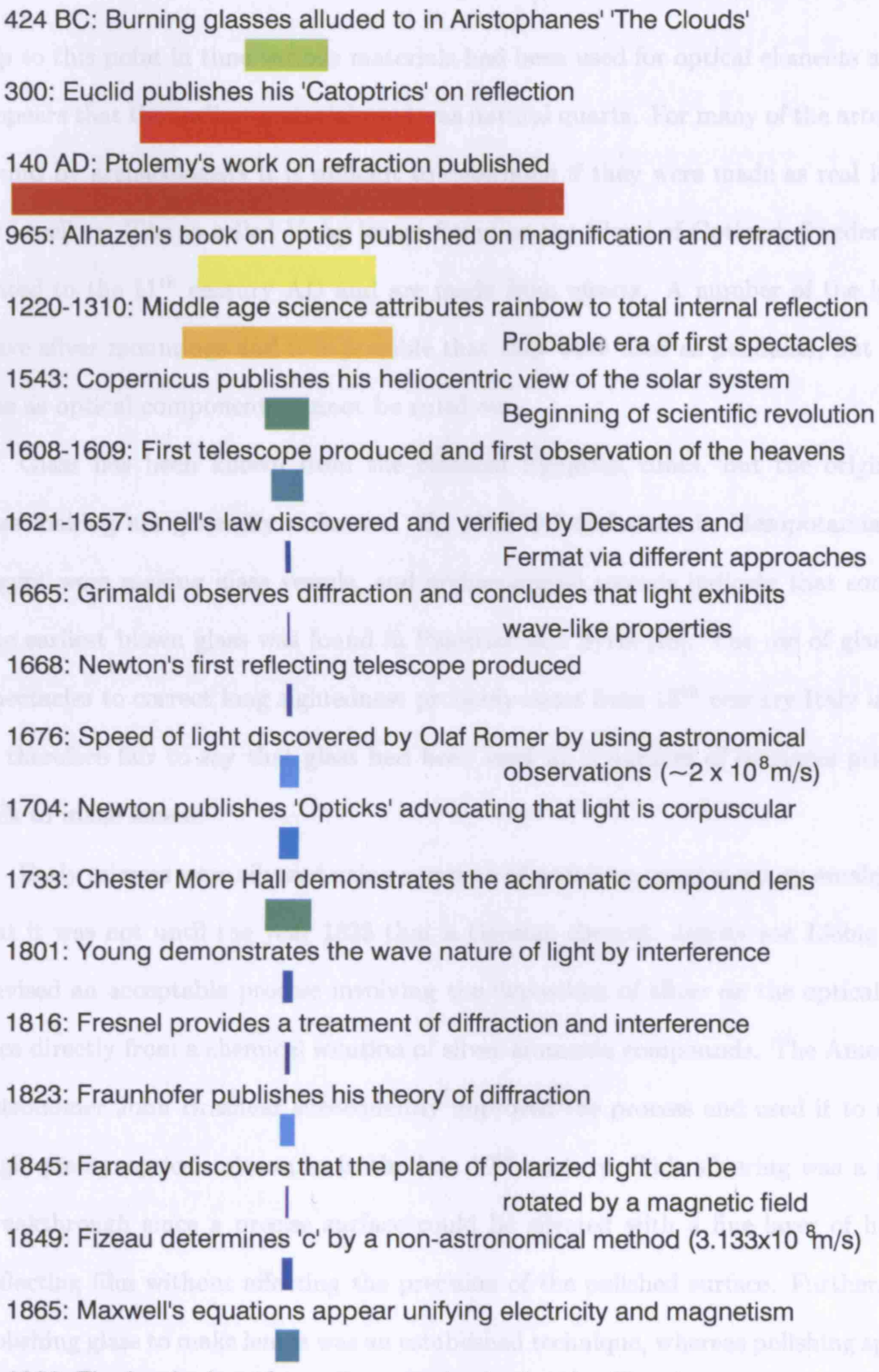


Figure 1.3: Optics timeline. See text for an explanation of the colour bars. Taken from [16, 2, 17].

1.2.5 Lenses, Mirrors and Telescopes

Up to this point in time various materials had been used for optical elements and it appears that the earliest material used was natural quartz. For many of the artefacts found by archaeologists it is difficult to determine if they were made as real lenses or jewellery. The so-called Visby lenses found on the island of Gotland, Sweden, are dated to the 11^{th} century AD and are made from quartz. A number of the lenses have silver mountings and it is possible that they were used as pendants, but their use as optical components cannot be ruled out.

Glass has been known from the classical Egyptian times, but the origins of glassmaking are generally unknown. By 1500 BC craftsmen in Mesopotamia and Egypt were making glass vessels, and archaeological records indicate that some of the earliest blown glass was found in Palestine and Syria [18]. The use of glass for spectacles to correct long sightedness probably dates from 13^{th} century Italy and it is therefore fair to say that glass had been used for a number of centuries prior to this to make lenses.

Early mirrors were silvered using a variety of processes using mercury amalgams, but it was not until the year 1835 that a German chemist, Justus von Liebig [19], devised an acceptable process involving the deposition of silver on the optical surface directly from a chemical solution of silver-ammonia compounds. The American astronomer John Brashear subsequently improved the process and used it to make high-quality optical telescopes in the late 19^{th} century. This silvering was a great breakthrough since a precise surface could be silvered with a fine layer of highly reflecting film without affecting the precision of the polished surface. Furthermore polishing glass to make lenses was an established technique, whereas polishing speculum was regarded as a black art and the surface finish of speculum tarnishes easily requiring re-polishing. Silvering can simply be dissolved away and re-applied to restore mirrors - a significant advantage over speculum mirrors.

In the science and astronomy world, the reign of big refractors for research gradually gave way to the reflector towards the beginning of the 20th century, when it was realised that producing and supporting larger lenses for new telescopes was too difficult and costly compared to producing and supporting reflecting mirrors of comparable performance. The earliest silver-on-glass telescope was built by C. A. von Steinheil in 1856 with a 4" aperture using Liebig's silvering process [20]. One year later, Foucault polished and silvered a 13" mirror for the Paris observatory.

Gradually, step by step, manufacturers were able to make larger and larger mirrors, usually by the more traditional polishing technique mentioned in section 1.2.6. Today, the largest monolithic mirrors are 8.4 m in diameter and two of them are used for the Large Binocular Telescope (LBT).

There is a common thread binding together the production of optical quality surfaces, whether made from speculum alloy or low expansion glass – it is the process of polishing. Without polishing, optical quality surfaces could simply not exist.

1.2.6 Classical Optical Polishing

Most optical quality surfaces have traditionally been polished using a lapping process. The surfaces to be polished are generally near-spherical and a full-size lap in the shape of the form to be polished is used in the presence of a polishing slurry such as a suspension of cerium oxide in water. Polishing can be considered as a ratio-metric or differential technique whereby doubling the work time doubles the material removal. This contrasts to cutting or diamond turning which is much more absolute.

For spheres, the spherical symmetry of the lap and of the part ensures that the tool conforms to the surface in all orientations and with any size of tool stroke. This allows a spherical surface to be polished with relative ease. This spherical symmetry is lost when an aspheric surface is produced. Here, a full sized tool cannot conform

to the surface except in one single position and this is exacerbated with off-axis aspheres. To mitigate against the effects of this non-compliance, sub-aperture tools have been used to produce aspherics which can conform approximately over a small part of the surface. The form is then controlled by changing the amount of time a tool is used over a particular part of the surface - the dwell time method, as used by several of the technologies mentioned in section 1.5.

Figure 1.4 shows a traditional pitch lap used to polish flats. The highly viscous pitch, the lapping surface, serves to hold the particles of cerium oxide which then modify the workpiece surface by a ductile (polishing) or brittle (grinding) process dependent upon the particle size and tool velocity. Lap substrates are commonly made from glass with similar properties to that of the workpiece such that their thermal expansion coefficients are matched. Figure 1.5 shows the pitch lap mounted on a legacy polishing machine at UCL acquired from the former Grubb-Parsons company. The lap would normally be mounted on the arms above the workpiece table where the arms are connected to eccentric drives to enable the lap's position and dwell time on the workpiece to be controlled.

Figure 1.4: A typical pitch lap for flats about 500 mm in diameter, courtesy of D. Brooks, UCL.

Polishing using machines such as that shown in figure 1.5 on page 35 still requires a skilled optician to produce the required surface form and the process involves an iterative procedure of polishing followed by measurement that can be extremely time

Figure 1.5: The UCL Grubb-Parsons 1 m class legacy polishing machine showing the eccentric drives and the polishing table.

consuming.

In recent years polishing technologies have been undergoing something of a revolution, partly due to advances in computers and computer control. Section 1.5.2 contains a summary of some of the recent polishing processes that have been developed over the last two decades. However, before summarizing the state-of-the-art, it is instructive to review some of the requirements for modern precision surfaces.

1.3 Design Drivers for Precision Surfaces

In the last two decades we have witnessed a dramatic change in the way that many optical components are designed and manufactured. Indeed, the art and craft of the traditional optician is being replaced by the application of science, engineering and computers. Now, the grinding and polishing processes are carried out using numerically controlled machine tool technology - technologies invariably developed within the last 20 to 30 years.

1.3.1 Spherical or Aspherical?

Optical designers have always understood that it is easier to create a spherical surface than an aspheric one with simple tools and age-old technology. The earliest lenses mentioned in this chapter were made using essentially the same methods as those used today. The only difference is that the tool is now often controlled by a machine rather than directly by the hand of the optical engineer.

Aspheric designs are desirable but have traditionally been associated with an order of magnitude increase in production costs at the very least. Their chief advantage is that they offer more degrees-of-freedom to the optical designer; a spherical surface has only a radius of curvature whereas an aspheric surface has a radius and also an asphericity which can consist of higher order terms. For example, a common way to specify an axially symmetric optical surface is to specify the sag, Z, at any point, h, from the optical axis [21, 22]:

$$Z = \frac{Ch^2}{1 + \sqrt{1 - \varepsilon C^2 h^2}} + A_4 h^4 + A_6 h^6 + \ldots + A_n h^n, \tag{1.1}$$

where C is the curvature ($= 1/R$), h is the distance from the optical axis, ε is the conic constant and the A_n are the aspheric coefficients. If the A_n are all zero and $\varepsilon = 1$, the equation reduces to the sag for a sphere.

The higher order terms in equation 1.1 can be used to balance out problems generated in other parts of an optical system. The increased number of degrees-of-freedom also mean that a given optical system can be designed with aspheric elements to give the same imaging performance as a spherical system with more elements. Off-axis sections and non-rotationally symmetric parts can also be specified enabling the use of unobstructed pupils or allowing the elements to be used as part of a mechanical surface. Therefore significant advantages can be gained in several areas by the specification of aspheric elements.

Modern polishing technologies have extended the envelope of optical production capability to allow these more complex forms to be produced with greater ease.

However, we must treat optical production as a complete process including design, production and metrology and this cannot be stressed too greatly if we are to succeed with some of the larger projects such as Extremely Large Telescopes (ELTs)[1], inertial fusion research optics and eventually fusion power stations. Without adequate metrology, it does not matter if we can polish parts with deterministic removal rates since we will not know where to polish. This is reflected in the well-known adage "*If you can't measure it, you can't make it.*"

Aspheric and more generally free-form surfaces present a significant difficulty to the metrologist, particularly the establishment of the base radius of curvature which is often an arbitrary radius with respect to where the departure is measured. It will be seen from the later sections in this chapter that metrology is currently a major limiting factor in the process cycle for producing several types of optical components.

1.3.2 Example Surfaces

There are numerous specifications in existence for the types of optical surfaces that may be required now and in the future as the following summarizes [23]:

- Large optics for astronomy, defence, space, high-power lasers and advanced instrumentation:
  - Size: 200 mm - 1 m today, 2 m and above for ELT applications.
  - Forms: Flat, concave and convex, spherical, aspheric, on- and off-axis, circular and non-circular.
  - Form-precision: frequently diffraction limited, but varies by application.
  - ELT segments require an absolute measurement of the radius of curvature.
- Large optics for extended UV photolithography:

[1]Current ELTs have primary mirrors in the range of 20–100 m in diameter.

  - Size: 300 – 600 mm, larger in future.
  - Forms: Mostly spherical, substrates are circular.
  - Form-precision: down to 5 nm P-V.

- Next-generation optics for astronomy, defence and instrumentation:
  - Size: 50 – 300 mm.
  - Forms: free-form surfaces and very large flats for laser projects.
  - Form-precision: target 100 nm P-V.

- Optics for consumer goods:
  - Size: typically 5 – 50 mm in diameter.
  - Forms: Spherical, on-axis aspheres.
  - Form-precision: 100 nm P-V.

- Prosthetic Joints:
  - Size: 75 mm.
  - Forms: Complex saddle forms (knees) and near-spherical forms (hips).
  - Form-precision: 1 μm.

- Industrial moulds and turbine blades:
  - Size: Typically up to 200 mm.
  - Forms: Complex and arbitrary free-forms often with very steep slopes.
  - Form-precision: approximately 1 μm.

The final two are examples of mechanical systems that require optical tolerances and can therefore benefit from optical manufacturing technology.

It is evident from this incomplete but illustrative list that many such optical, and non-optical, components of the future will involve tight form constraints on

non-spherical forms and that more free-form surfaces will be employed, particularly in the defence sector where mechanically conforming surfaces might be required. Such free-form surfaces can be specified in many ways but a common way is to use Non-Uniform Rational B-Splines (NURBS) which provide a convenient and unified mathematical basis for representing analytical and free-form entities. This means that they can describe precisely any conic section and can be used to describe free-form surfaces of any shape and complexity in a form that is now becoming an industry standard [24].

As a specific example of the kinds of optical surfaces required at the present time and in the near future, the following section contains an outline of the proposals for some of the current Extremely Large Telescope projects.

1.3.3 Optical Telescopes - The Future

Astronomy has been described by Rowan-Robinson [25] as *'the beacon of the physical sciences, inspiring wonder among people from all walks of life'*. This has been evident from the time of Galileo when humans first observed the heavens using optical aids. Many early telescopes were built by the users themselves, including Galileo, but slowly, enterprise and public use led to the establishment of observatories, both state and private. The eventual establishment of university observatories led to the active research programmes in astronomy and to the increased requirement for large telescope apertures.

From the early days of telescope manufacture alluded to in section 1.2.5 to the present day there has been a continuing trend to design and make ever larger telescopes. The desire for larger telescopes stems from the desire to gather light from fainter and fainter sources led by fascinating science goals. Soon, the current set of 8 – 10 m class telescopes will be superseded by much larger instruments and several designs exist with apertures in the range of 20 – 100 m known as *Extremely Large*

Telescopes (ELTs)[2].

To date several ELT projects exist with apertures greater than 20 m. In the US we have the TMT (Thirty Metre Telescope), a Caltech project, and also the GMT (Giant Magellan Telescope) led by a corporation formed between several universities and research organizations.

The TMT design specifies a segmented primary mirror with 492 segments each 1.45 m across corners - the same dimension as the ESO ELT design's segments [27]. The telescope is a Ritchey-Chretien design with an adaptive convex secondary mirror. The basic statistics for the TMT are shown in table 1.1.

Primary mirror diameter	30 m
Radius of Curvature at PM vertex	60 m
Conic constant	-1.000953

Table 1.1: TMT primary mirror basic configuration. The conic constant of -1.000953 indicates that the primary mirror is a hyperboloid.

The GMT telescope is a stopgap design, based on seven segments each with a circular outline and a diameter of 8.4 m [28]. The telescope is a Gregorian type with a hole in the central segment. The six outer segments are each identical in form simplifying the construction - only two types of segments need be made, namely the centre one and six copies of the outer segment design. Table 1.2 shows the basic statistics of this telescope's primary mirror. The 8.4 m GMT segments are currently being manufactured at the Steward Observatory mirror lab in Arizona. The segments are light-weighted monoliths as distinct from all other ELT projects

[2]One metric of telescope performance is the resolving power and the Rayleigh [26] formula states that objects that subtend an angle of $1.22\lambda/d$ can just be resolved, where d is the telescope aperture and λ is the wavelength of light studied. Therefore, smaller angles can be resolved by choosing larger apertures.

Equivalent primary mirror diameter	25.4 m
Radius of Curvature at PM vertex	36 m
Conic constant	-0.99829

Table 1.2: GMT primary mirror basic configuration.

where the segments are much smaller solid hexagonal forms which are comparatively thin for their diameter. The maximum form error on each GMT segment is of the order of 40 nm P-V for spatial wavelengths less than 0.1 m. The optical test of the segments is briefly mentioned in section 1.4 on page 46.

There are two competing designs for ELTs in Europe, ESOs Overwhelmingly Large Telescope (OWL), a 100 m diameter design, since downsized to 42 m and renamed the E-ELT, and the EURO50 project led by Lund University for a 50 m diameter telescope. The optical specifications of these telescopes are being driven by the science goals, but a thorough survey of the science goals is beyond the scope of this thesis. However, a tabular summary of the model used by Hook and Gilmore [29] in the European ELT case study serves to illustrate the science drivers which are broken down into three main categories:

- Detection of extra-solar planets and the study of circumstellar environments
- Stellar population and galaxy formation studies
- Studies of high-redshift environments

Table 1.3 summarizes the results of the European ELT case study carried out over 4 years from 2001 [29]. These science goals are a very exciting but challenging pursuit. They are driving the optical specifications in all aspects. For example, extra-solar planet detection will involve imaging a star and examining the region around the main star for a smaller object, possibly many tens of orders of magnitude fainter.

Aperture	(Exo)Planets and Stars	Stars and Galaxies	Galaxies and Cosmology
20 m	• Direct detection of Jovian-mass planets in wide orbits around nearby solar-type stars • Radial velocity search on fainter stars (increasing available volume by a factor of 200)	• Resolved oldest stellar populations in Magellanic Clouds and Local Group dwarf spheroidals • Resolution of the brightest giant stars in galaxies in the Virgo cluster • Observations of halo giants in Local Group galaxies (high-res spectroscopy)	• Ly-alpha emission-line spectroscopy from 6<z<10 • Possible detection of z~10 objects (depending on their nature)
30 m	• Imaging of young (<10Myr) Jovian planets around stars in star-forming regions up to 75pc away • Detection and classification of mature Jovian planets around stars within 10-20pc • Possible detection of one Earth-like planet within ~5pc	• Age/metallicity measurements of resolved populations in M31/M32 at ~750kpc (imaging) • Determination of star formation and chemical enrichment histories of galaxies out to Cen A (nearest active galaxy)	• Possible detection of z~10 objects (depending on their nature) • Spectroscopy of "earliest galaxies" found by JWST • IGM studies to z~10 using brightest GRBs as background sources
100 m	• Survey of 1000 solar-like stars and direct detection of Earths within 30pc • Time-resolved photometry of Earth-like planets (albedo & weather) • Spectroscopy of earth-like planets and search for "Biomarkers" • Study of entire exo-planetary systems	• Age/metallicity measurements of resolved populations in M87 (in Virgo cluster at 16Mpc) • Detailed study of galaxy formation in a representative sample of the Universe	• Detection of z>10 objects • Spectroscopy of "galaxies" to z~20 (depending on their nature). Such objects may even be resolved with a 100m • IGM studies at z>10 (GRBs, QSOs, PopIII SNe as background)

Table 1.3: A summary of the key ELT science drivers as a function of the aperture size, from [29].

This requires excellent stray light rejection which drives the specification for the surface texture, form control and edge form since the mirrors of such ELTs are segmented and badly finished edges could be a source of much stray light.

The optical components of ELTs lend themselves to the idea of using a complete and unified optical production system and a design study was undertaken for the EURO50 project, a 50 m diameter Gregorian design, commissioned by the Instituto de Astrofisica de Canarias [30]. It stressed the need for a systematic approach to polishing and metrology to ensure that segments are matched to tolerances. This matching introduces the concept of traceable metrology to optical manufacture – referencing the measurements to the international definition of the metre as defined by the International Bureau of Weights and Measures - and forms an important part of the work on the NPL/UCL swing-arm profilometer.

Traceable metrology is important since the mirrors for ELT projects will take several years to manufacture and during this time, different metrology devices may be used in the production processes. Traceable metrology simply allows measurements to be made where each measurement result can be traced back to the definition of the metre via an unbroken chain of measurements, each with a stated uncertainty. Provided that the overall uncertainty in the measurement (of radius, for example) is less than the maximum permissible error, then each segment can be made to specification. Without traceable metrology, the segments could all be manufactured with slightly different radii depending on the metrology tools in use at the time of manufacture. With traceable measurements, we can be sure that the instruments are giving the same answers and therefore that the segment's radii are all matched to within tolerance provided that the measurement uncertainty is low enough.

To date most ELT designs are in the 20–100 m diameter range and some project proposals are still at the evolutionary stage in terms of overall design. The European Southern Observatory's (ESO) design for the Overwhelmingly Large Telescope

(OWL) had the largest aperture at 100 m diameter, giving the proposed telescope a collecting area greater than that of all the current telescopes today [31]. The telescope was regarded as too costly and ambitious to manufacture and has since been down-sized to give a primary mirror of 42 m diameter.

The two proposals reviewed here, the EURO50 design and ESO's E-ELT, both utilize aspheric primary and secondary mirrors and are very similar in their general specification. The EURO50 telescope seems to have remained at the design stage at the time of writing.

The EURO 50 project Gregorian telescope design has a 50 m diameter primary mirror (f/0.84) comprising some 618 hexagonal off-axis aspheric mirror segments, each of 2 m in diameter across the flats. The aspheric departures from the best-fit sphere of these segments ranges from 3 μm for the innermost ring of segments to 300 μm for the edge segments. Furthermore, the base radii of curvature of the individual segments must match each other to within 0.5 mm to ensure that the required performance goals are met. The tolerance on the form of the segments is very strict at 36 nm RMS which equates to approximately $\lambda/20$ at 632.8 nm.

ESO's ELT, now consists of a 42 m diameter primary mirror and two candidate optical system designs [32]. One design calls for a Gregorian system with a concave secondary, whilst the favoured design giving near-perfect imaging quality across the entire field of view is composed of 5 mirrors with a convex secondary. In either case, the f/1 aspheric primary mirror will be 42 m in diameter and composed of some 1146 hexagonal segments each of 1.45 m corner-to-corner. Here, the design documents specify an RMS form error of less than 50 nm at 632.8 nm.

The concept designs for the ESO E-ELT and EURO50 telescopes are shown in figure 1.6 and the corresponding specifications are indicated in tables 1.4 and 1.5 respectively.

For both designs one thing is very clear from the specification - the metrology

Figure 1.6: Left: ESO ELT, 42 m primary 5 mirror telescope. Right: EURO50 ELT, 50 m primary Gregorian telescope. Both show background vehicles to illustrate the scale.

Surface	RoC (m)	Diameter (m)	Form	Conic constant (ε)
M1	-84	42	elliptical	-0.999356
M2	6.386	3.4	elliptical	-0.777912
Greg Foc	3.36	0.28	spherical	0
M4	64	2.8	hyperboloidal	-1.362129
M5	61	2.4	elliptical	-0.722396
Nas foc	3.34	0.29	spherical	0

Table 1.4: The ESO E-ELT 5-mirror optical configuration at the time of writing. RoC is radius of curvature.

Surface	RoC (m)	Diameter (m)	Form	Conic constant (ε)
M1	-85.000	50.64	elliptical	-0.999358
M2	6.5739	4.10	elliptical	-0.777798
B1	∞	1.3	flat	0
Image surface	3.456	0.76	spherical	0

Table 1.5: EURO50 optical design for the aplanatic Gregorian configuration at the time of writing. RoC is radius of curvature.

goals are very challenging not only due to the requirements for form accuracy and radius of curvature matching, which is ±0.5 mm for Euro50, but also from the point of view that the metrology has to work reliably over the complete production life-cycle of the segments. With several hundred segments this could equate to several years. Indeed, at the final stage of polishing a large, high-quality mirror, 90% of the total process time may be taken up entirely by handling and metrology [33]. Testing is possible using conventional technology, but often, time and cost become critical and limiting factors for the success of such projects.

1.4 Stitching and future ELT projects

The proposed test for the GMT segments can be compared to the principle of a stitching test to illustrate the differences in procedures. The GMT segments will be tested by an interferometric test employing several corrective elements in the optical path between the interferometer and part [34]. The corrective elements form a null corrector – two spherical fold mirrors and a CGH that together transform the interferometer's wavefront into the correct test wavefront. Most of the aspheric compensation is made by the first tilted fold mirror whose diameter and radius of curvature are 3.75 m and 25.5 m respectively. The second mirror serves to further correct the wavefront and shorten the optical path, and is 0.77 m in diameter and has a radius of curvature of 1.37 m. A CGH corrects the remaining aberrations in the system. The large fold mirror is placed at the top of a 28 m test tower and its position must be controlled and known very accurately in order to eliminate introducing unwanted aberrations into the test. Positional accuracy must also be maintained for the other elements in the test, but an allowance is made for some error since the segments themselves can be moved and stressed once mounted in the telescope. The test is a technical *tour de force* incorporating many novel features, but it is also incredibly complex and requires weeks of work to set-up and use,

not to mention the cost of the tower and null optics. The test is then subject to turbulence, vibration and air stratification problems. Furthermore, using a modern interferometer with 1000 pixels across a diameter of the measurement artefact gives a lateral resolution of 8.4 mm per pixel - this effectively filters out features with linear dimensions of less than about 20 mm.

A test of the segments could be made, in principle, by stitching data from an interferometer or other wavefront sensor. If a wavefront sensor can be mounted on a suitable computer controlled multi-axis stage, measurements can be made with much shorter path lengths using a weakly diverging test lens. The part would then be tested with virtual confocal points with the interferometer located a few mm from the surface under test as shown in figure 7.1 on page 286. The interferometer would be translated over the surface and measurements would be made at predefined positions computed from the part's aspheric departure. The stitching algorithm would then process the data to compute the surface form error.

The advantage of this principle is that the lateral resolution of of the test is much higher - 1000 pixels across, say a 150 mm aperture - which is good news for the measurement of the higher spatial frequencies that can be introduced by sub-aperture polishing techniques.

The main requirement for a stitching test is that the aspheric surface deviation of the part under test over the diameter of one sub-aperture must not introduce too much slope change for the interferometer or wavefront sensor. In interferometers high slopes cause high fringe densities, and at some point sub-Nyquist sampling will occur if the slopes are too great. This sets a practical limit on the sub-aperture size and therefore slopes must be evaluated on a case-by-case basis for the parts under consideration.

A sub-aperture test can therefore limit the optical path length and minimize errors caused by the vibrations of large test towers. They can also provide increased

lateral resolution. This all comes at the expense of having to design a stitching system to perform the metrology and computation. However, it should be noted that a single CNC stitching system could be designed to test a range of different segment geometries - another advantage over the test tower scheme where the test tower is designed around the optical surface required. We will return to these - and other - metrology issues in later sections following on from the next section reviewing polishing technologies.

1.5 Polishing Technologies: a brief review

The study of optical polishing and grinding techniques has quite recently resulted in a large number of scientific publications concerning new approaches to the production of precision optical components. It is therefore useful to summarize the conventional wisdom and that produced more recently as explained in the following sections.

1.5.1 Classical Polishing

Classical polishing is still carried out by amateur telescope makers and some major scientific research and production establishments. The traditional techniques, mentioned in section 1.2.6 of making pitch laps, full- and sub-aperture tools are used along with the skill and experience of the craft optician.

1.5.2 Computer Controlled Polishing

Several computer controlled polishing processes of recent development have been demonstrated each using a novel approach to the issue of tool conformance. Some key technologies are summarized in the subsections below.

Stressed Mirror Polishing

Lubliner and Nelson [35, 36] demonstrated stressed mirror polishing in 1980. They successfully demonstrated a technique that involves applying a uniform pressure across the back of a mirror blank and applying shear forces and torques to its circumference. A sphere is then polished into the stressed blank and then the forces are removed allowing the mirror to spring into the desired shape. Much of the content in the literature deals with the method of computing the forces required to produce an off-axis paraboloid as a prototype for a 10 m ground based telescope[3]. The technique is applicable to blanks that are not extremely thick and also to blanks that are moderate aspheres. For thicker mirrors and more extreme aspheres the forces involved become too high for many of the optical glasses currently in use.

ITEK Optical Surfacing

A computer controlled polishing process was recently developed at the Litton Itek corporation by Zimmerman et al. [37]. The process, using Computer Numeric Control (CNC) of a sub-aperture lap has produced off-axis aspherics with claimed surface figures of 20 nm and convergence values[4] of 0.9 in the late stages of polishing. The material removal is computed by the convolution of the tool work function with the path of the tool over the surface. The process does not require rotation of the substrate and can thus be used to produce non-circular off-axis segments.

Stressed Lap Polishing

Work in the 1990s at the Steward Observatory mirror lab [38] involved the production of a 1.8 m f/1 ellipsoid. Martin et al. polished a sphere into the 1.8 m blank with a rigid sub-diameter lap before using a stressed lap to polish the final ellipsoid.

[3]UC Ten-metre telescope.

[4]Convergence value: $C = \frac{\delta_n}{\delta_{n-1}}$, d_n = RMS surface figure error of the nth cycle, d_{n-1} = RMS surface figure error of the n-1th cycle.

In their work, the lap consisted of a 600 mm diameter aluminium disc whose shape could be continuously changed using 12 actuators. Their rationale was to retain the advantages of large rigid tools by making them able to conform actively to the surface being polished, no matter what the position on the mirror surface. Recently the lap was used successfully to polish the 8.4 m primaries for the Large Binocular Telescope.

Reactive Atomic Plasma

Developed at the Lawrence Livermore laboratory by Verma et al., reactive atomic plasma polishing [39] is a sub-aperture chemical process well suited to silicon carbide whereby material removal is accomplished by using an atmospheric-pressure gas plasma containing reactive fluorine. The process uses inductively coupled radio-frequency power to excite an argon and CF_4 plasma at atmospheric pressure. Material from the substrate is removed by the reaction of CF_4 with the SiC, where the SiC is converted to SiF_4 and escapes as a gas. A key advantage of the method is that it does not exert any force on the optic being polished. Recent work by Fanara et al. [40] has demonstrated this process on ULE glass. Volume removal rates of 0.5 mm^3/s have been demonstrated, but heating and therefore thermal expansion becomes a problem with higher removal rates.

Ion Figuring

Ion figuring uses a directed neutralized ion beam, see for example, [41], to remove material by physical bombardment and is therefore a sub-aperture process. The surface to be figured is placed in a vacuum chamber where a CNC translation system moves the beam over the surface to be figured. The algorithms usually proceed by assuming that the material removal is isotropic, linear and proportional to the dwell time. The dwell times are computed by consideration of the beam removal function, the optic surface error map, the required optic form and the tool/beam

path. It is interesting to note that the final form correction of the Keck telescope mirror segments was carried out using an ion figuring process by the Eastman Kodak company after they were found to be in error after the initial stressed mirror polishing process.

Precessions

Over the last seven years, work at the author's laboratory, OSL/UCL, in partnership with the spin-out company Zeeko Ltd, has resulted in an automated process for producing many types of optical surfaces [42, 43]. The process, called *Zeeko classic*, involves the use of a sub-aperture sized tool operating in the presence of a conventional polishing slurry. The tool used is an inflated rubber membrane with a spherical form covered in a flexible polishing surface. The tool is spun at up to 1500 rpm and applied to the surface at an angle to the tool axis of rotation. The tool is also precessed around the local surface normal.

There are several advantages to this method of polishing. The membrane moulds itself to the local aspheric surface form ensuring good contact everywhere and the spot size can be varied by either increasing the load on the tool or increasing the inflation pressure. Precession of the tool is used to give excellent texture to the polished surface and results in a near Gaussian polishing spot or influence function. The process uses the dwell time method to remove material in a similar manner to other sub-aperture processes. Figure 1.7 shows the latest incarnation of the Zeeko/UCL polishing machine, the Intelligent Robotic Polisher for optics up to 1200 mm in diameter undergoing final testing. The IRP 1200 also has the ability to polish parts using a fluid jet called Zeeko Jet using technology licensed from TNO-TPD in the Netherlands.

Figure 1.7: The Zeeko IRP 1200 CNC polishing machine employing the Zeeko Precessions process for parts up to 1200 mm diameter.

1.5.3 Modern Use of Polishing

Many of the recent processes are well suited to the production of aspheric forms since they are mostly sub-aperture processes and use tools that comply with the local surface. Many processes, particularly Precessions, RAP and Fluid Jet can produce very repeatable material removal functions which can be used in conjunction with a surface error map to perform very high precision form correction. However, some care must be taken in choosing the size of the removal function and the tool path otherwise mid-spatial frequency forms may be generated on the surface. These result in a lower Strehl ratio and therefore limit the performance of the optical surface.

1.6 Metrology and Instrumentation

The recent progress in polishing technology has been accompanied by some advances in form and surface metrology. The amount of desktop computing power available cheaply has to some extent helped this with new metrology tools delivering a very

high density of data points from high speed data acquisition systems. Traditional optical tests designed to measure spheres or conics and have been well documented in for example, Malacara [21]. The mainstay instruments have been commercial Fizeau and Twyman-Green interferometers and various strategies can be employed to allow them to test different forms, as summarized in the coming sections.

However, even considering the advances in the speed of data acquisition or the density of data points, one thing is clear: *accurate 3-D surface metrology is still a challenge where large apertures, steep aspheres and free-form parts are concerned.*

For example, a large aspheric lens element could be measured by a linear profilometer where several diametrical scans are made over the part. The profilometer would require a vertical measurement range greater than the maximum sag of the part under test which could be many tens of mm. To construct a contact probe with an accuracy less than 100 nm with the required measurement range is a difficult task - one is not currently available on the market, nor in research labs. A work-around for this is to move the probe over a circular datum that matches the base radius of the part under test such that the probe now only has to measure the aspheric departure of the part. We have now transformed the problem of making an accurate long-range probe into the problem of moving an accurate short-range probe - which can be bought commercially - over the required base sphere with a very high level of precision ($\simeq$30 nm max error). This second problem is also very challenging and no such device exists on the market to achieve this. However, the NPL/UCL swing-arm profilometer attempts to solve this problem as described in chapter 5.

Another continuing challenge is the interferometric measurement of large diameter aspheric mirrors and lenses for optical systems as exemplified by the current crop of ELT projects such as the segments for ESO's E-ELT. The outermost segments have an aspheric departure of 1.2 mm - equivalent to almost 2000 waves at 633 nm. Tests being devised for these segments are very complicated involving many

sub-optical systems and verification procedures. Furthermore, the optical tests are restricted to the design segments. Whilst the principles of the test can be re-used for different segments, the equipment, including test towers, fold surfaces and any holograms must be redesigned and manufactured for a new set of differing segments.

Free-form surfaces of all sizes also pose the same problems except here we do not always have a convenient reference datum, for example, we may wish to produce a precision saddle where there is positive and negative curvature at the stationary point. Therefore most difficult task for freeform surfaces is to produce the required measurement datum whether optical or physical. This is especially difficult when the parts deviate from simple conics - examples include jet engine turbine blades and head-up display projection lenses.

We may well ask: what is a free-form surface? Recent work by Walker et al.[44] has defined a free-form surface as: *'a surface that deviates significantly from a rotationally-symmetric form'*. Parts that are off-axis sections of rotationally-symmetric aspheric parents, such as off-axis paraboloids, hyperboloids etc. were also included in this definition as special cases.

We now turn to a review of some of the metrology instruments currently in use to measure optical form, and some of the common problems encountered with their use.

1.6.1 Contacting Metrology Devices

Coordinate Measuring Machines

Coordinate Measuring Machines (CMMs) have long been used in engineering and manufacturing to provide traceable measurements of machine parts. The usual CMMs have 3 orthogonal axes of motion in a Cartesian system coupled to a mechanical probing system such as a touch trigger probe[5]. The probe is made to

[5]A touch trigger probe sends an electrical signal to the CMM controller to read the coordinate axis position encoders once a pre-set force has been applied to the probe. In this way, consistent

contact the surface under test at predetermined points to suit the measurement strategy. Reference is made to a datum point and the axis slideways where each axis of motion generally contains a position encoder measuring the position relative to the datum point. Therefore before use, a calibrated CMM must be reset to the datum position before measurements can be made.

Full calibration of a CMM is a lengthy process, particularly when traceable metrology is required. This problem is further exacerbated by the length of time taken to make a large number of measurements on a surface - measurements are usually taken point by point with the measurement tip removed from the object during translation along the axes.

Figure 1.8 shows a Fischers Aktien-Gesellschaft (FAG) CMM at the National Physical Laboratory (NPL) that was later converted into the NPL/UCL swing-arm profilometer. Details of the mechanics, operation and author's contributions to the swing-arm can be found in chapter 5.

Measurement of parts with linear dimensions larger than 1 m becomes difficult with a CMM since the size of the structural members must increase to maintain dimensional stability on the sub-micron level. Constructing such a large CMM is therefore an expensive and challenging task, not to mention the time required to acquire sufficient data points to map the surface under test at the resolution required.

Mechanical Profilometer

Traditional profilometers work by moving a measuring probe along a single axis and make measurements with respect to a datum-feature such as a line or a surface, which can be mechanically or optically defined. Ruby tipped probes are commonly scanned along the surface under test and by monitoring the probe's tip position the surface form is elucidated as a function of the linear position along the surface and of

readings of the surface form can be achieved.

Figure 1.8: The FAG CMM before conversion to a swing-arm profilometer. The probe shown is an example of a touch trigger probe that can register contact in two orthogonal directions.

the form of the datum. The main drawback of such devices is that they only usually have one axis of motion and the probe speeds must be kept to below a few mm per second to avoid losing contact with the surface. Calibration of profilometers is similar to the calibration of CMMs - the tip shape must be well known or significant errors can result when measuring on steep slopes.

Commercial profilometers are available with trace lengths up to 300 mm and resolutions down to the nanometre level. As with CMMs, large profilometers are difficult to make if measurements are required on the nanometre scale over traces above 1 m in length. However, profilometers are particularly useful for measuring ground parts as the precision required is less than that for polished parts, say 0.1–1 μm. A profilometer for such parts was constructed at UCL by Yang using a granite beam as the mechanical stage for a linear air-bearing, with an optically defined datum consisting of a laser beam in free air [45]. Measurement is performed by a precision linear encoder probe with a resolution of 25 nm as described in

[46]. It is intended that this instrument will be used to measure ground parts at the Ultra Precision Surfaces (UPS) facility of UCL and Cranfield in North Wales, complementing the NPL/UCL swing-arm profilometer.

1.6.2 Non-contacting Metrology Devices

Interferometry

Many different optical interferometers have been used in optical testing including the ubiquitous Fizeau and Twyman-Green types which are the most popular and well developed commercially. The operating principles are similar for most interferometers, regardless of the mechanical arrangement. A reference wavefront, forming the measurement datum, is mixed with the wavefront reflected or transmitted from the object under test. The error between the datum and the wavefront to be measured is encoded in the interference pattern produced.

Modern optical interferometers are capable of measuring flat, spherical and near-spherical surfaces much to better than $\lambda/20$ P-V at 632.8 nm. Modern systems can make high resolution measurements in a matter of seconds once the mechanics of the test have been set up. Most issues occur when attempting to test longer focal length optics where vibration or air turbulence become problematic, and when testing steeper aspheres where the fringe density becomes too high without the use of a null compensator. Large aperture forms also pose significant difficulties, particularly convex parts, as illustrated in figure 6.14 on page 241.

Two main techniques are used in optical interferometry to infer the shape of a wavefront [21]. The first is phase mixing where a known reference wavefront is mixed (interfered) with the unknown wavefront. The difference between the two is encoded in the interference pattern. The second technique is that of self-referencing where the unknown wavefront is divided into two beams. One beam is then modified by a known amount (e.g., using lateral shear) before it is recombined with the unmodified

wavefront datum. Again, the aberrations are encoded into the resulting interference pattern.

Fizeau Interferometery

A Fizeau interferometer utilizes the principle of Newton's fringes and phase mixing where the fringes are formed between a reference artefact and the surface under test [47]. Typically, the reference artefacts will be flats or spheres, but aspheres can be made at a cost. However, the forms of such aspheres are generally as difficult to verify as the parts required themselves.

Interferometers of this kind are readily available in apertures from a few mm up to 300 mm (4–12"), anything much larger becomes incredibly costly to manufacture, with 6" systems costing ca. £100,000.00. Twyman-Green interferometers, although different in mechanical construction, require similar data-reduction techniques.

The original Fizeau interferometers were simple **static phase interferometers** since the optical test was set up and the resulting fringe pattern was simply photographed or recorded by a CCD before analysis. In the static fringe method, the fringe contours are measured and the coordinates of the fringe maxima and minima are entered into a computer for the phase unwrapping. A polynomial surface is then fit to the unwrapped data from which the surface statistics are computed. With modern static phase systems the phase is recovered from the interferogram modulo 2π before phase unwrapping is used to calculate the surface form error. The main problem is that non-uniform illumination can introduce significant errors and the algorithms are only able to deal with simple fringe patterns.

Phase-shifting is a technique used in optical interferometry to remove non-linear effects from affecting the measurement result, and involves taking at least 4 times as much phase data compared to static phase types. Such non-linear effects can arise from the non-uniformity of the interferometer's illumination and variabilities in

gain across the CCD camera. Furthermore, static phase interferometry algorithms usually require between 10 and 30 fringes of tilt to be present in the interferogram as a spatial carrier. This tilt introduces re-trace errors into the measurement and serves to increase the uncertainty (section 4.11). Phase-shifting interferometers work best with nulled fringes – another significant advantage.

Phase-shifting is normally accomplished by moving the part and the interferometer relative to each other by up to one wavelength using a piezoelectric (PZT) phase-shifter. In most phase-shifting Fizeau interferometers the transmission element is shifted by the PZT relative to the interferometer and the surface under test.

Phase-shifting interferometers provide a convenient means of inferring the optical phase at any point coming from the surface under test, and several methods exist for inferring the phase (modulo 2π) from a series of interferograms with a known phase-shift between them. The process is most easily understood by considering the acquisition of four interferograms where there is a physical phase difference of $\pi/2$ between each one.

Let two beam interference occur in a Fizeau interferometer. We let the test arm signals be subscripted by, a, and the reference arm signals by, b. It can be easily shown by considering the electric fields of the two interfering beams that an interferogram contains the following terms from a standpoint of intensity:

$$I(x,y) = I_a(x,y) + I_b(x,y) + 2\sqrt{I_a(x,y)I_b(x,y)}\cos\phi, \tag{1.2}$$

where, x and y are the pixel coordinates in the interferogram, the I are the intensities in the beams and δ is the optical path difference between the two beams at that point - the quantity required which is a function of the shape of the surface under test.

To perform the 4-step algorithm, we record 4 interferograms but we deliberately change the phase by $\pi/2$ between each measurement. We thus have 4 interferograms

(omitting the x- and y-indices):

$$I_1 = I_a + I_b + 2\sqrt{I_a I_b}cos(\phi + 0) \quad (1.3)$$

$$I_2 = I_a + I_b + 2\sqrt{I_a I_b}cos(\phi + \pi/2) \quad (1.4)$$

$$I_3 = I_a + I_b + 2\sqrt{I_a I_b}cos(\phi + \pi) \quad (1.5)$$

$$I_4 = I_a + I_b + 2\sqrt{I_a I_b}cos(\phi + 3\pi/2). \quad (1.6)$$

We convert the cosines into sines and have the following:

$$I_1 = I_a + I_b + 2\sqrt{I_a I_b}cos(\phi) \quad (1.7)$$

$$I_2 = I_a + I_b - 2\sqrt{I_a I_b}sin(\phi) \quad (1.8)$$

$$I_3 = I_a + I_b - 2\sqrt{I_a I_b}cos(\phi) \quad (1.9)$$

$$I_4 = I_a + I_b + 2\sqrt{I_a I_b}sin(\phi). \quad (1.10)$$

Now we can eliminate the unknowns, namely I_a and I_b, to arrive at an expression for the phase, ϕ, by evaluating:

$$\frac{I_4 - I_2}{I_1 - I_3} = \frac{\sin\phi}{\cos\phi} = \tan\phi. \quad (1.11)$$

We can therefore evaluate the phase directly from the 4 interferograms to produce an answer modulo 2π. This modulo 2π phase map must be integrated (unwrapped) to obtain the actual phase by counting the 2π phase steps along the integration path.

It will be evident that using phase-shifting interferometery eliminates problems associated with non-linear illumination, since each point on the surface under test is measured 'independently of the others' and with high precision. The 4-step technique can be extended to allow a large number of interferograms to be captured and analysed with any convenient value of phase step. Phase-shifting therefore gains significant advantages over conventional static fringe interferometry.

The main drawback of this technique is that the phase-stepping takes a finite amount of time - the time taken to step the reference surface 4 times, plus the time

taken for the elements to settle down, and plus the time taken to take four exposures using the CCD camera. For the Wyko 6000 system, a common system with a fairly typical operating characteristic, the acquisition time is approximately 167 ms. In this time vibrations and turbulence between the test and reference surfaces may affect the quality of the measurement result.

This type of optical interferometry is useful for flats, spheres and very mild aspheres with low aspheric departures[6]. When high departures are considered, the fringe density becomes too high to perform phase-recovery. A Computer Generated Hologram (CGH) [48, 21] placed after the transmission sphere can be used in these circumstances to test the part. The hologram is a binary representation of the fringe pattern that would result if the ideal aspheric wavefront required were interfered with a tilted plane wavefront.

In use, the CGH is placed in the interferometer test set-up where the CGH and interference fringes produced by the interference of the reference wavefront and the aspheric surface under test produce a moiré pattern encoding the difference between the CGH and the interference fringes. Here, the main problem is in verifying the hologram's pattern errors and the hologram's alignment.

The perils of long optical path lengths have already been mentioned, with the main errors coming from turbulence, vibration and the change in the refractive index of air with temperature in the presence of the inevitable thermal gradients - not such a problem for the newer vibration insensitive interferometry. Large aperture parts with long focal ratios (ELT segments) are prime candidates for such long path lengths. Testing larger parts inevitably means that spatial resolution is compro-

[6]The definition of 'low' here is difficult to quantify since in involves several factors such as the CCD pixel size, the interferometer magnification and the shape of the parts being tested. Often manufacturers simply quote the maximum slope difference that can be measured under the best circumstances.

mised[7] and in the case of the GMT segments the 8.7 mm lateral resolution means that features whose sizes are less than about 20 mm will not be seen on the part.

A schematic diagram for a commercial Fizeau interferometer is shown in figure 1.9, showing the main components of a typical modern instrument as shown in figure 1.11 on page 63. A key advantage of Fizeau types over Twyman-Green[8] types is the common path of the light inside the interferometer which makes them less sensitive to vibration. However, vibration of the surface under test relative to the interferometer mainframe still remains an issue for conventional phase-shifting and static fringe types.

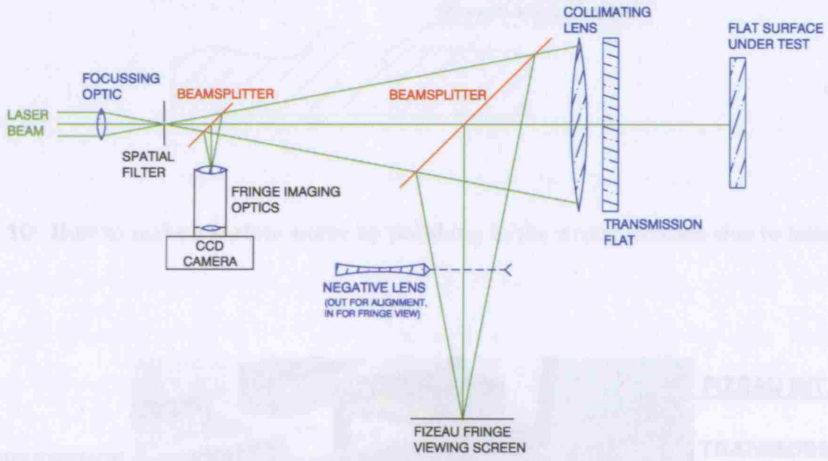


Figure 1.9: Schematic diagram of a commercial Fizeau interferometer redrawn from [47].

A key aspect of using Fizeau interferometers is when the absolute form of the surface under test (SUT) is required. Traditionally, in a craft sense, an interferometer would be used to compute the Peak-to-Valley surface deviation and the RMS surface deviation of the part. The part would then be polished by hand with reference to the

[7] A typical modern Fizeau interferometer will have about 1000 pixels across its field of view.

[8] A schematic for a commercial Twyman-Green type is shown on figure 2.8 on page 106.

surface error map and so the process would be iterated until the required form was reached. Computer control of the polishing now means that the phasemaps must be correctly scaled in the lateral dimensions such that the polishing machine can be programmed to polish the correct areas on the optic. This is especially important when small tooling sizes are used since an error of a few mm can mean that we polish the wrong part on the surface and we may make the surface worse as indicated in figure 1.10.

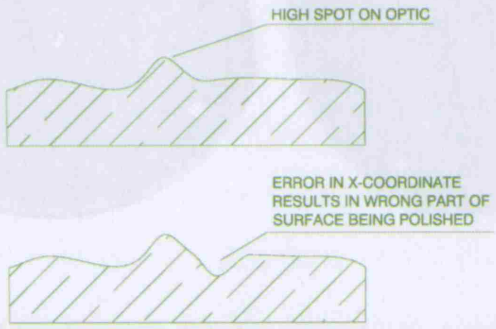


Figure 1.10: How to make a surface worse by polishing in the wrong location due to lateral scaling errors.

Figure 1.11: A typical modern Fizeau phase-shifting interferometer testing a plano part. The part under test is a 305×155 mm fused silica flat that has been oiled to a larger flat to eliminate reflection from the back surface.

Figure 1.12 on page 64 shows two overlapping sub-aperture measurements of an

f/1 sphere of 200 mm diameter. The figure clearly shows the distortion introduced into the projection of the interferograms on the interferometer CCD. This distortion is dealt with in section 4.1.1 and must be corrected if errors like those shown in figure 1.10 are to be avoided.

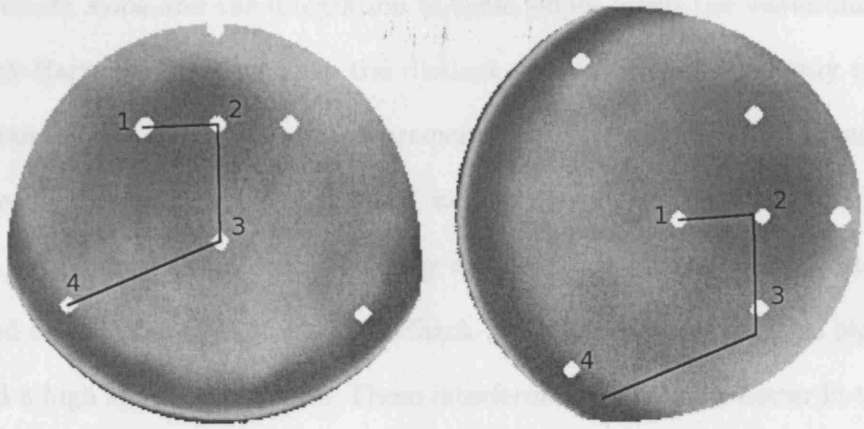


Figure 1.12: Effect of interferometer distortion (aspect ratio correction and scaling about the optical axis required) on phasemap. A different portion of the part was measured in each case, with a small common overlap, with a 90 degree rotation and shift of optical axis between measurements. The phasemaps were then oriented such that the part appears 'the right way up' here. Note how the numbered fiducial points do not match between the left and right phasemaps. Black line shows the fiducial relationships for the left-hand image.

Shack-Hartmann

Shack-Hartmann (S-H) wavefront sensors (WFS) are commonly used in Adaptive Optics (AO) and for wavefront sensing. The principle is the same as that for the Hartmann test[21], but instead of impinging onto a full-aperture plate with an array of holes, the wavefront at the pupil plane is instead re-imaged onto an array of microlenses [49]. The array of lenslets serves to focus the incoming wavefront down to a series of spots, each corresponding to a particular lenslet. If the incoming wavefront is perfectly plane, a regular array of spots is formed on the image plane. If the wavefront is aberrated then the local slope of the wavefront at each lenslet

will be different and the resulting spot from a given lenslet will be displaced from the null position. Reconstruction of the incoming wavefront is achieved by assuming a continuous (smooth) wavefront and by computing the local tip/tilts required to place the spots in the configuration measured. The S-H sensor therefore measures the wavefront slope and the integration of these slopes yields the wavefront phase.

Shack-Hartmann sensors have the distinct advantage that they only require a single frame of data to make a measurement and they can therefore be used with high-speed CCD cameras to make fast measurements for Adaptive Optics (AO) or form measurement purposes. Recently simultaneous phase interferometers have appeared on the market that can rival Shack-Hartmann sensors giving a high frame rate and a high spatial resolution. These interferometers are also better in the sense that they sample the returned wavefront from the whole of the SUT, rather than just at the points covered by the lenslet array in the S-H sensor. However, S-H sensors can be made to have a greater dynamic range than interferometers and to therefore measure steeper aspheres depending upon the test required.

Curvature Sensing

Curvature wavefront sensors establish the form of an incoming wavefront by considering the intensity distribution on two planes symmetrically placed about the instrument entrance pupil. The second derivative of the wavefront phase, the wavefront curvature has been shown by Roddier et al. [50, 51] to be proportional to the difference in intensity passing the two planes about the instrument entrance pupil - an approximation to the axial intensity gradient. If I_1 is the intensity on one plane, and I_2 is the intensity on the other then Roddier showed that:

$$\frac{I_1 - I_2}{I_1 + I_2} = -k\nabla^2\phi, \tag{1.12}$$

where k is a proportionality constant and ϕ is the wavefront phase at the instrument entrance pupil. Solving this Poisson equation gives the phase of the incoming wave-

front. If the intensity on the two planes is recorded simultaneously then curvature sensing can be used as a wavefront sensor, and this type of sensor is known as a defocus curvature sensor as the two planes contain defocused images of the pupil.

Roddier et al. were the first to use the particular technique of curvature sensing to provide in-situ testing of telescope optics based on work in their 1988 papers[loc cit.]. The problem with curvature sensing is that computation of the wavefront phase assumes a continuous function of curvature - in the same way that Shack-Hartmann testing assumes that slopes are continuous. This causes problems with the integration when unusual apertures are met such as those with central obscurations.

Recent work by Greenaway et al. [52, 53, 54, 55, 56] has attempted to generalize on the work of Roddier et al. by introducing diffractive elements to perform the simultaneous imaging onto a single camera, and to introduce aberrations other than defocus into the system. This work was of importance to this project and will be dealt with in section 2.2.1.

1.7 Measuring ELT Segments

ELT segments require testing at various stages in the production process. Firstly as rough ground forms during and after grinding. Here, the form error will be between approximately 1 – 50 μm P-V and there are a number of established techniques available such as multilateration, CMM metrology and custom profilometric techniques that can provide <1 μm of precision at each measurement point. Alternatively, long-wavelength interferometry at 10.6μm can be used provided that the aspheric departures are manageable. The most challenging metrology occurs at the polishing stage.

The initial polishing of the ELT segments, taking the form error down from a few μm to 0.5 – 1 μm P-V, will allow a variety of optical measurement techniques to be used, including interferometry at 633 nm. With form accuracy requirements of 30

–50 nm RMS over 2 m, the segments are technically challenging; however, the long radius of curvature of over 80 m (E-ELT and EURO50) means that they are quite flat, even the segments with the most aspheric departure, with central sag values of the order of 3–6 mm. This makes traditional interferometry difficult due to the long path lengths and air turbulence, but this is an advantage for mechanical techniques such as profilometry since the probe is always near-normal to the surface[9]. A swing-arm profilometer (SAP), such as the NPL/UCL SAP described in chapter 5 maintains the contact probe normal to the surface of the best-fit sphere, whatever the base radius of curvature. However, with profilometric techniques, the accuracy is generally limited because the geometry of the profilometer must be fully characterised to eliminate straightness and angular errors in the profilometer's slideways and bearings.

A feasible solution to testing ELT segments is to use CGHs [57], which can achieve an accuracy of 10 nm with spherical profiles [58]. However, the main drawback is that long air-paths are still encountered and that a new CGH is required for each segment type. Furthermore, assessing the uncertainties associated with alignment and the dimensions on the CGH remain a concern.

The authors of the EURO50 design study report concluded that interferometric techniques are feasible, but that a better strategy may be to combine several technologies such as collecting data simultaneously from point-by-point and by areal methods. This would give the precision of the point-by-point method combined with the coverage and collection rate of the areal methods. If the two techniques could be physically related then this may establish a route to the traceability of the optical measurements.

[9]For a small deviation from perpendicular the cosine error is small.

1.8 Project Outline and Aims

The example of ELT segments mentioned in the previous sections represents the state of the art in regard to extremity of quality, complexity and form. The ability to perform high-quality quantitative measurements with the necessary precision, speed and dynamic range for an affordable and feasible manufacturing process does not exist today.

One approach to testing ELT segments, mentioned in section 1.7 is to combine technologies, both hardware and software, in hybrid designs. This thesis addresses one strand of work designed to investigate combining the advantages of novel wavefront sensing and data stitching with the future aim of incorporating them onto a machine tool. The techniques of metrology data stitching are therefore explored, involving testing small regions of a whole surface using an optical wavefront sensor and joining or stitching the regions together in software.

This stitching has been facilitated by the novel approach of placing removable physical fiducial marks on the surfaces being tested, and by using these fiducial marks to determine the transformations required to stitch the metrology data.

The primary deliverable is a set of techniques and Matlab software routines to perform the stitching. Several ancillary techniques have also been developed to aid the main stitching algorithms and these are primarily concerned with data acquisition, locating the fiducials in the metrology data, known as *fiducial centroiding*, and the efficient use and storage of data within the main software algorithms.

In parallel to this work, UCL have been working with the National Physical Laboratory (NPL) on the design and construction of a swing-arm profilometer described briefly in chapter 5 and more comprehensively by Efstathiou [59].

1.8.1 OMAM Project

The work forms part of the DTI (EPSRC/PPARC) Optical Manipulation and Metrology (**OMAM**) project led by Heriot-Watt university. The project was designed to exploit novel wavefront sensing technologies for use in optical metrology projects by the project partners. A key deliverable of the project was to be a Phase Diversity Wavefront Sensor (PDWFS) by Heriot-Watt, for use by the project partners including UCL, NPL, Zeeko Ltd, Scalar Technologies and BAE Systems.

Schedule slippage and software problems at Heriot-Watt meant that a sensor was never delivered despite repeated promises to the contrary. A review of the proposed PDWFS and how an *important work-around was sought and implemented* can be found in sections 2.2.1 and 2.3 respectively.

In particular, the UCL strand of the OMAM project was designed to investigate

(A) Large optics measurement by swing-arm profilometry. The swing-arm profilometer also provides a host platform for a wavefront sensor and/or the more usual contact probe to give a hybrid system.

(B) Optics measurement by stitching metrology data.

(C) Ancillary techniques and technologies to facilitate (A) and (B) above.

(Here, 'large' is taken to mean up to 1 m in diameter)

Therefore the work in this thesis serves to fulfil the requirements of (B) above.

1.9 The Problem: what is being attempted?

Let two portions of an optical surface be measured at different points with some common overlap area between them, and let the measurement data be of the form:

$$\mathbf{W} = \begin{bmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix}, \quad \&, \quad \mathbf{W}' = \begin{bmatrix} x'_1 & y'_1 & z'_1 \\ \vdots & \vdots & \vdots \\ x'_n & y'_n & z'_n \end{bmatrix}, \tag{1.13}$$

where each row in $\mathbf{W}$ and $\mathbf{W}'$ represents a measurement point.

We define stitching as finding a geometrical transformation to be applied to $\mathbf{W}'$ such that the points in the overlap area, $\mathbf{W} \cap \mathbf{W}'$ are *matched* in the best possible manner. When the transformation that matches the points and gives the least error has been applied, the measurements are said to be *stitched*.

The outlined text describes stitching two sub-apertures together and this idea can be extended to accommodate any number and shape, provided that there is overlap between any two adjacent sub-apertures.

Let the data be measured on a machine with perfect positioning and readout systems (i.e. no errors or noise present and infinite resolution). We can derive the stitching transformation by reference to the mechanical machine state at each measurement position provided that the machine readout is sufficiently precise. Otherwise, we can compute the stitching transformation by using some common points of interest between overlapping measurement areas.

In this work, a hybrid approach is considered: we use the machine state, or optical set-up configuration information to generate an initial and approximate transformation and then use the approach of common points to refine the stitching.

The novel aspect of using physical fiducials placed on the surface to test is used to establish a set of common points between two overlapping sub-apertures.

1.10 Applications of Metrology Data Stitching

Numerous problems can be solved by stitching metrology data. The simplest is where the aperture or focal ratio of an optical test system is insufficient to image the whole surface of the object under test – this is the classical sub-aperture problem. This situation may arise in the testing of large flats and convex or concave spheres and unless a full aperture test can be contrived, the only solution is to perform a sub-aperture test.

Stitching may also be used in the case of steep aspheres. A single measurement of the whole aperture may be impossible due to the maximum change of slope on the part across the aperture. However, over a more localised region the slope change may be small enough that it may be measurable by a conventional method such as interferometery. Here, sub-aperture tests can be carefully arranged such that the whole surface can be measured by stitching.

Sub-aperture stitching can also be used to increase the spatial resolution of a measurement by using beam compressors to provide a given resolution over a smaller area.

Applications are not limited to optical surfaces; potential applications include measuring industrial moulds, prosthetic knee joints, and indeed profilometry data as touched upon in chapter 5 of this work. Stitching can therefore be used anywhere where patterns of points need to be matched.

Hurley and Cattell used a novel approach to matching groups of points in psychometric tests [60] now known as Procrustes analysis which uses the singular value decomposition to determine the stitching or matching transformation. A Procrustes algorithm is used in later chapters as a special case of the stitching problem to stitch two sub-apertures together - a scheme that can be extended when there is an array of $n \times 1$ sub-apertures to be stitched together.

1.11 Existing Stitching Technologies: A Brief Review

Several stitching techniques and algorithms have been reported in the past and whilst many have been in the public domain, proprietary algorithms also exist. Most algorithms share the same goal: to deduce a set of geometric transformations that will stitch together several sub-aperture measurements with the minimum residual error. The condition of minimizing the error in stitching has often led to the use of least-squares techniques in the final determination of the transformation parameters. What follows is a very brief review of the main technologies, in chronological order, that are of direct relevance to this work.

There exist sub-aperture testing approaches that do not involve stitching. Jensen et al. [61] contains a review of these which generally seek to reduce a series of sub-aperture interferograms (phasemaps) to produce a full-aperture map. The techniques of Chow et al. [62] and Thunen et al. [63] involved non-overlapping sub-apertures and are not considered as stitching here. Both works involve sub-aperture configurations as shown in figure 1.13. In this scheme, it is not possible to align

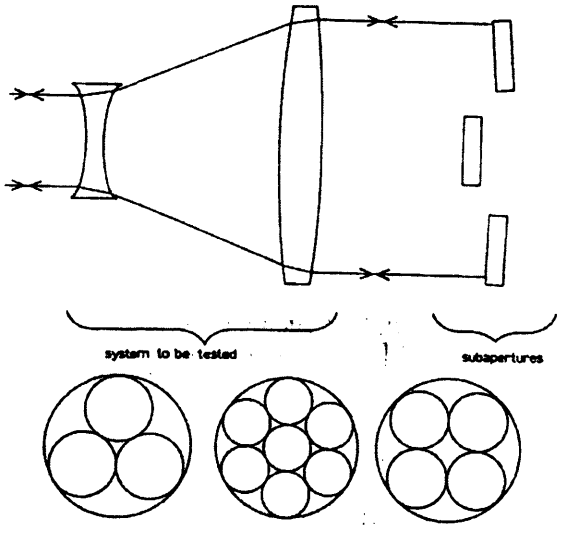


Figure 1.13: Sub-aperture test configurations as used in [62] and taken from [61].

the sub-apertures to remove the relative piston and tilt terms, nor is it possible to

have prior knowledge of these quantities. Consequently, this uncertainty results in a reduction in the test accuracy.

In 1991 Chen et al [64] described a Multi-Aperture Overlap Scanning Technique (MAOST) to overcome the shortcomings of the previous sub-aperture testing methods by making each sub-aperture test overlap with its neighbours. In this way, the amount of tip and piston between sub-apertures can be calculated. Here, a wavefront (sub-aperture measurement) can be expressed as

$$W(x,y) = A + Bx + Cy + W_0(x,y), \tag{1.14}$$

where $W_0(x,y)$ is the wavefront measurement without tip, tilt and piston, and A, B and C are the piston, tip and tilt coefficients respectively. The connection (stitching) of the points in the two sub-apertures $W_1 = W_1(x,y)$ and $W_2 = W_2(x,y)$ proceeds via the simple matrix equation:

$$W_1(x_i, y_i, \phi_i, 1) = W_2(x_i, y_i, \phi_i, 1) \cdot \begin{bmatrix} 1 & 0 & \Delta B & 0 \\ 0 & 1 & \Delta C & 0 \\ 0 & 0 & 1 & 0 \\ p_x & p_y & \Delta A & 1 \end{bmatrix}, \tag{1.15}$$

$$\text{or :} \quad W_1(x_i, y_i, \phi_i, 1) = W_2(x_i, y_i, \phi_i, 1) \cdot \mathbf{T}. \tag{1.16}$$

Here, ΔA, ΔB and ΔC are the relative piston, tip and tilt respectively. The p_x and p_y are the displacement transformation coefficients which are known from the translation introduced. After measurement, the wavefronts are written in terms of orthogonal polynomials and least-squares fitting is used over the overlap area. Chen [loc cit.] quotes $\lambda/600$ RMS in the literature as an achievable precision, although the value of λ is not quoted and this is believed to be for simulated data. The work also introduced the concept of the slightly misleading term, the aperture magnification factor, defined as the ratio of the diameter of the full aperture to that of the sub-aperture. This has since been renamed the extension factor by Fleig et al [65].

One drawback of this approach is that the orthogonal basis polynomials must be tailored due to the irregularly shaped overlap areas that are not identical from overlap to overlap region. Furthermore, relative, or polar, rotation between sub-apertures is not considered in MAOST.

A technique to perform obtain a full-aperture measurement of a large flat by sub-aperture stitching was proposed by Otsubo et al. [1] by means of a similar procedure. Here, the flat to be measured was translated under a Fizeau interferometer such that a series of overlapping sub-aperture measurements was made. The work considered the tip/tilt and piston between adjacent sub-apertures and figure 1.14 indicates the geometry of the stitching for just two sub-apertures.

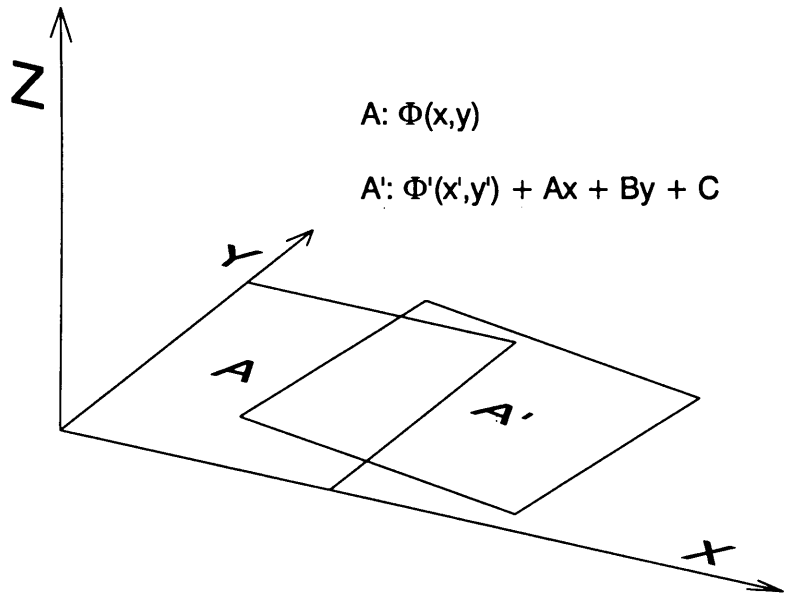


Figure 1.14: Stitching geometry redrawn from Otsubo et al. [1]. A is the reference sub-aperture to which all others are stitched. A' is a sub-aperture to be stitched to A by altering its tip/tilt and piston coefficients.

Referring to figure 1.14, $\Phi'(x', y')$ is stitched onto $\Phi(x, y)$ by the following transformation,

$$\Phi(x, y) = \Phi'(x - x_0,\ y - y_0) + ax + by + c, \tag{1.17}$$

where a, b and c are the tip, tilt and piston terms respectively. x_0 and y_0 are the

origin shifts between A and A'. Again, a least-squares method is used to find the coefficients a, b and c i.e.,

$$\text{minimize} \quad \sum_{A \cap A'} \{\Phi(x,y) - [\Phi'(x_0 - x_0,\ y - y_0) + ax + by + c]\}^2. \tag{1.18}$$

Summation occurs at all points in the overlap area. The methodology is developed to allow stitching on $n \times m$ sub-apertures such that the global least-squares condition is met over all sub-apertures. In this work, as for the MAOST work, polar rotation between sub-apertures was not considered.

Bray [66, 67, 68, 69] used the approach of Otsubo et al. in a commercialised system to measure large flats for laser work. A precision x-y stage was used to servo a Fizeau interferometer over the large parts to the sub-aperture positions. After measurement, each sub-aperture was then perturbed using tip, tilt and piston until the least-squares condition was met.

In 1996, Sullivan et al. [70] sought to verify a mechanical roundness measurement of a Zerodur sphere by making 36 overlapping measurements of the sphere using an interferometer and rotating the sphere by 10 degrees about the vertical between measurements. The row of data representing the equatorial region was then stripped out from each phasemap and used as a profile. Thus 36 profiles were stitched with the requirement that each profile match its neighbour in position, tip and power in a least-squares sense. Figure 1.15 indicates the type of profiles and the overall measurement result compared to an independent measurement of the sphere by the Physikalisch-Technische Bundesanstalt (PTB), Germany. The least-squares problem looks daunting as there are 36 equations to be solved with almost as many unknowns. However with modern computing power and high-level languages, such problems can be formulated and solved in a few hours of work, with the actual process of number crunching taking only a few minutes of CPU time.

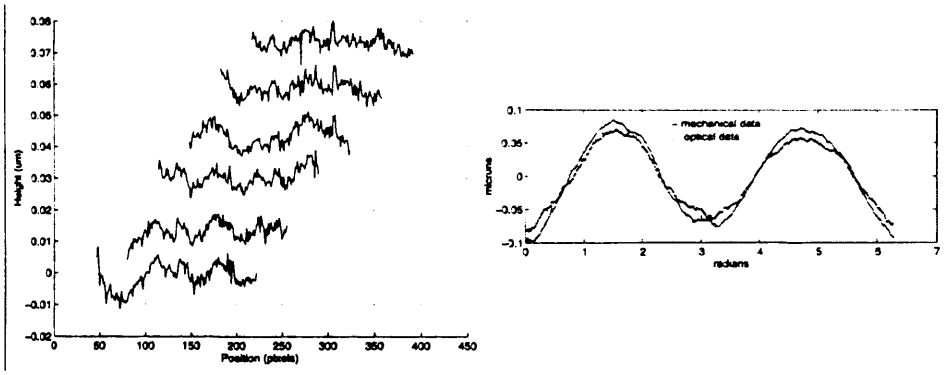


Figure 1.15: Overlapping profiles from [70] stitched using a least-squares method.

1.11.1 QED Stitching Interferometer

The QED Spherical Stitching Interferometer is an instrument developed at Rochester NY for the measurement of high numerical aperture spheres by stitching together sub-aperture interferometry data. The instrument consists of a commercial Fizeau interferometer and a precision 6-axis motion system. In operation, the part is rotated about the confocal point and after selection of an appropriate transmission sphere the testing configurations for the 6-axis system are computed as depicted in figure 1.16. The surface is divided up into regions corresponding to the footprint of the test wavefront and these are so arranged such that there is an overlap between adjacent testing positions. This is shown schematically in figure 1.16. The motion system then moves to each position required and waits for the interferometer to measure the form of each sub-aperture before moving on to the next position.

At each measurement point, the surface form is recorded along with the configuration of the 6-axis system. Stitching of the sub-apertures is then performed as a least-squares optimization using the knowledge of the 6-axis system's co-ordinates at each measurement point. At the time of writing, the device has been used with some success on convex spheres of varying diameters. The flowchart taken from [71] shown in figure 1.16 serves to illustrate the measurement procedure. The details of the algorithm are the intellectual property of QED Technologies, but it is gen-

erally understood that the stitching proceeds via a least-squares algorithm taking into account the aberrations introduced between the sub-aperture measurements, including tip, tilt and piston. Furthermore, the literature indicates that the errors in the reference wave are subtracted from the results. The test indicated in the

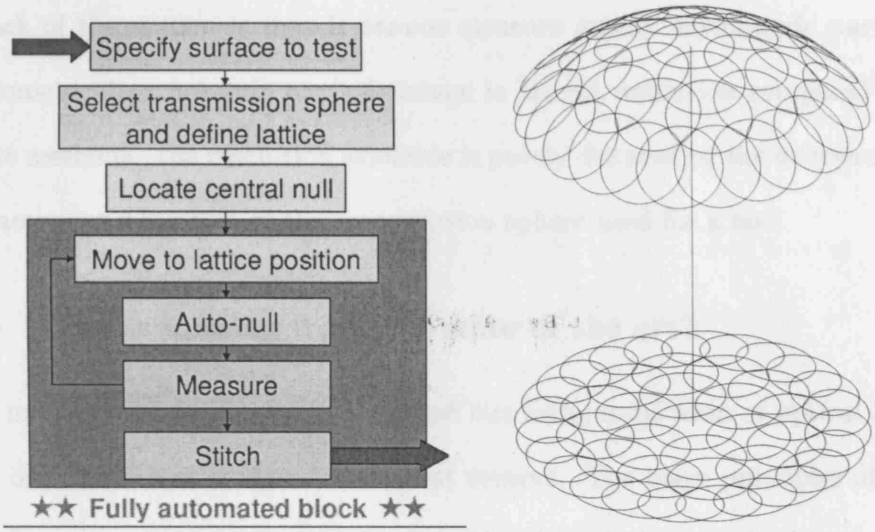


Figure 1.16: Left: Flow-chart for the QED stitching interferometer. Right: Concept of the lattice map showing the intersections of the interferometer's test beams upon the surface under test and their projection on the x-y plane.

right of figure 1.16 has an extension factor[10] of approximately 4.13, and various configurations of overlapping sub-apertures are possible giving a different overlap area. Indeed, in their work, Tricard et al. [72] have reported using sufficient overlap such that each point on the surface is measured in at least two sub-apertures. The QED technology is currently the only automated commercial stitching device in the marketplace.

1.11.2 Zygo Verifire Asphere

The Zygo Verifire asphere is a recent example of an asphere metrology system designed to measure axially symmetric parts up to 130 mm in diameter with a maxi-

[10]Defined on page 73 as the ratio of the diameter of the full-aperture to that of the sub-aperture.

mum aspheric departure of 10 μm. The instrument incorporates a Fizeau interferometer and a displacement interferometer system and measures annular zones on the test part. Each zone is measured whilst the distance along the z-axis is noted before joining them together to produce the final measurement result. The main drawback of the system is that it cannot measure axially asymmetric parts. The interferometer does not gain any advantage in lateral resolution compared to sub-aperture methods - the resolution available is purely dictated by the interferometer's CCD camera and the f/# of the transmission sphere used for a test.

1.11.3 What's missing from the state of the art?

Today, most optical form testing is carried out using some form of optical interferometer or more recently optical wavefront sensors. The basic principles of optical testing have remained unchanged for many years and thus the basic components have also remained unchanged. It is the ancillary equipment that has seen the most change with interferometers being equipped with adds-on and software for phase-shifting interferometry and now for vibration insensitive instantaneous phase-shift interferometry. With the advent of high resolution lithographic systems it is now also possible to produce high quality CGHs for generating aspheric wavefronts, yet the traditional problem of testing such lenses and gratings does remain a problem however.

The state of the art in optical specification requires the advent of new testing methods and the current limitations for testing large, long (and short) focal length surfaces are:

- The optical path length becomes too long for a viable measurement, in cost and in practice. It is not practical or feasible to perform an 84 m long test. Auxiliary optics are required to shorten this to less than 30 m.

- The time taken to perform the measurement becomes a significant or dominant

factor. The acquisition times for interferometric measurements may only be measured in ms, but the setting up and verifying an optical test can take days.

- A precise determination of the basic parameters is often required: radius of curvature, aspheric departure, form error, clear aperture and surface texture. For example, the EURO50 specification requires that segment's radii are known to ±0.5 mm, and that the P-V form error is less than 38 nm from the base asphere. Interferometric measurement does not measure radius of curvature and so a procedure to do this is required. This might involve making a master sphere and performing differential measurements to compute the parts' radii.

- Moving parts between polishing or grinding machines and the metrology tool is hazardous. For example, if a part weighing 80 kg, the approximate weight of an ESO segment, is placed even gently on edge, extremely high forces are generated due to the high mass and small contact areas. This can be disastrous for an ELT segment causing the ultimate demise of a given segment.

- Fixing parts is difficult, especially when parts have to be removed and replaced back in the same position. During a form correction and measurement iteration, the part will need to be placed back in the metrology position such that its datum (often the back-surface) is in the same location as it was during the previous metrology cycle. This positional accuracy often needs to be on the level of 10-50 μ m [73].

- Full aperture testing of some parts may be difficult, especially very fast or slow parts.

- Today's instruments have a limited slope measurement capability. Larger slopes can be measured at longer wavelengths but at the expense of the measurement precision. For example, we might test at 10.6 μm, but here we may only achieve a resolution of 41 nm with an 8 bit digitization, as opposed to 2.4

nm at a wavelength of 633 nm.

- Satisfactory methods to test optical surfaces of large dimension, large deviation from sphericity or planarity and surfaces of high numerical aperture are not commonly available, as has been previously explained.

Walker has advocated the method of combining technologies as mentioned in section 1.7 on page 1.7 [30]. Interferometry cannot always give a radius of curvature measurement, but this may be possible with swing-arm profilometry, see chapter 5, which also provides a direct measure of a part's aspheric departure. Furthermore, swing-arm profilometry can introduce possibility of traceable metrology to ensure that, for example, ELT mirror segments are matched in radius of curvature.

The effects of long optical paths cannot be overstressed. Controlling vibration, removing the effects of turbulence and correcting for refractive index changes in air over lengths up to 85 m is very problematic. In vertical columns of air up to 10 m in length, laser beams have been known to deviate by several mm from the straight line path [74].

A recurring use for stitching metrology measurements is that they can be used to avoid the long path lengths associated with testing large aperture optics with long focal ratios. This is currently one of the main problems for ELT segment metrology and at present the only means of mitigating against long paths is to use auxiliary optics to fold and shorten the beam length. However this strategy increases the number of surfaces in the test beam that require verification. Simply testing at the centre of the base sphere, as described by equation 1.1, is not possible as often the aspheric departure is far too high.

Stitching optical metrology data can avoid long path lengths by using easy-to-make small-aperture diverging transmission elements to make surface measurements over small portions of a larger surface with a virtual confocal point. Here, the interferometer may be located just a few mm above the optical surface under test

and the virtual confocal point may be several tens of metres above this. The slower interferometer reference surfaces are easier to construct and measure than faster ones since for ELT segments, they will, by definition, be almost flat with central sags of less than 1 mm. This could form the basis of a significant reduction in the measurement error budget - a 0.1% change in refractive index of air over 85 m is 8.5 mm difference in optical path length (OPL), this reduces to 100 μ m OPL at 100 mm.

Finally, the hazards of moving parts cannot be overstated. A study by the late Dr Richard Bingham in the 1990s revealed that a significant number of optics were damaged in some way during manufacture [75]. It would be far better to test and polish optical surfaces on a single platform than to move them between instruments.

This thesis attempts to alleviate these problems of testing optical components by employing sub-aperture stitching to establish the basis of an extensible system that may be employed in future as part of an on-machine metrology system, to eliminate some of the problems discussed in this chapter, and to allow polishing and metrology to be performed on the same hardware platform.

1.12 The Remaining Chapters

The stitching problem and the approach taken will be established in the next chapter including the significance of physical fiducial marks to the metrology instruments used. The techniques developed and used for determining the positions of the fiducials are covered comprising an assessment of the requirements for an automated machine - after all, we wish to apply the principles in a CNC production environment in the future.

Chapter 3 deals with the development of a coherent mathematical system for the transformations required to stitch datasets. The foundling stitching algorithm, considered as a series of block operations, is built up and since the chapter deals

mainly with matrix operations, the Singular Value Decomposition is then considered for the special case of stitching an $n \times 1$ array of sub-apertures together.

In chapter 4 the stitching algorithm is developed and numerical optimization is encountered. A Matlab optimization program is applied to this particular problem to give a general stitching solution to the $n \times m$ sub-aperture stitching problem considered. As an extra, the algorithm for the special case of stitching $n \times 1$ sub-apertures is built up complementary to the main algorithm for the $n \times m$ case.

Chapter 5 presents an outline of the NPL/UCL swing-arm profilometer and how it can be linked to this work. Some of the author's contributions to the profilometer are also summarized.

Chapter 6 deals with stitching some simulated and real data and the procedures required in transforming the data from the metrology device.

Chapter 7, the closing chapter, contains the conclusions and a critique of the work and will be found useful in the further development of the stitching algorithms.

CHAPTER 2

The Approach, Instruments and Fiducials

"There ain't no rules around here! We're trying to accomplish something!"

– Thomas A. Edison, U. S. Inventor (1847–1931)

2.1 The Stitching Problem and a Possible Approach

Before considering candidate technologies for the physical metrology process, the stitching problem and an outline of a possible approach needs to be identified. In chapter 1, several pre-existing stitching algorithms have been summarized, either for general metrology or for more specific applications. Often these algorithms are proprietary, where exact details are not available, or even patented where full details are available to the public domain but the techniques cannot be copied or reproduced.

In this work a new approach is outlined with the aim of generating new intellectual property for commercial exploitation by the project partners. This means that existing work cannot be used, and a new technique to stitch data must be found. As

the story unfolds, it will be seen that the surface fiducialization provides a means of establishing the stitching points in the data, and that a numerical optimization is used to find the best stitching parameters.

One sub-aperture measurement will consist of measuring the physical form of a surface using any of the techniques of optical metrology; in particular wavefront sensing and interferometry. Thus:

> The idea of sub-aperture stitching in this work is simply stated as *taking the series of overlapping sub-aperture measurements of a pre-fiducialized surface, and using the fiducial data from the overlapping areas to compute the stitching transformation* to form an effective full-aperture measurement.

Fleig [65], has succinctly defined stitching as 'Synthesizing a full-aperture phase-map from multiple sub-aperture maps with a set of mathematical techniques'.

In this work, the *prior knowledge* in the overlap regions will consist of the points occupied in the data by the physical fiducial marks placed on the surface before measurement. It is therefore a requirement that these points can be recognised in the output data of the measurement instrument.

Reflective vs Transmissive testing

It is useful to mention that this work focuses on measuring surfaces in reflection and not in transmission. A surface's form will be measured by the retro-reflection of a known test wavefront and the subsequent analysis of this reflected wavefront. Stitching could be applied to testing parts in transmission, such as testing the homogeneity of a piece of glass, or testing a lens. When testing in transmission the measurement result is the sum of the errors in the part's surfaces and the errors from the inhomogeneity of the glass making up the part - the variation in refractive index within the part. Therefore, to completely map a part in transmission

requires measurements of all the transmitting surface forms and also a measurement of the homogeneity once the surface forms are known. Measuring in single reflection is much simpler - we directly obtain the difference between the reference wave and the test-surface reflected wave. All work in this thesis is motivated by the needs of testing in reflection, although the procedures could be applied to testing in transmission.

2.1.1 Simple Stitching

In the simplest stitching case, we imagine two overlapping data sets each measured in its own coordinate frame on a particular metrology device that provides data in x, y, and z coordinates. Let the instrument be displaced by an amount (x_0, y_0) between measurements and let there be no other aberrations or noise present in the system. Let the two data sets be represented by W and W', in the form,

$$\mathbf{W} = \begin{bmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \\ z_1 & \dots & z_n \end{bmatrix}, \quad \mathbf{W}' = \begin{bmatrix} x'_1 & \dots & x'_n \\ y'_1 & \dots & y'_n \\ z'_1 & \dots & z'_n \end{bmatrix} \tag{2.1}$$

where the data may represent an absolute surface form or a deviation from a datum-surface such as the test wavefront. Furthermore, in equations 2.1 to 2.9, the term 'points' may be taken to mean either the actual data points or to only the points considered/nominated as fiducials – the analysis is the same in either case. The stitching is performed by the translation

$$\mathbf{W}'_{\text{new}} = \mathbf{W}' + \begin{pmatrix} x_0 \\ y_0 \\ 0 \end{pmatrix}, \tag{2.2}$$

i.e. adding the translation vector on the right of the equation to each point represented by a column in $\mathbf{W}'$, and by forming a larger array $\mathbf{W}_F$ containing all the points from $\mathbf{W}$ and $\mathbf{W}'_{\text{new}}$. If the stitching is exact then the points in the overlap

area will be duplicated and one set may be discarded (or we may take the mean value at each point). Otherwise some form of filtering and regridding has to be carried out on the whole array $\mathbf{W}_F$.

Fleig has used the extra points in the overlap area to establish a likely error for the stitching [65]. Here, there will be n values for the phase at each pixel in an overlap area, where n is the number of overlapping sub-apertures at that point. The standard deviation of these values for each pixel can be considered as the residual noise following stitching. In the best circumstances, the noise should be spatially uncorrelated, and have a magnitude that approaches the raw phase measurement noise value divided by the square root of the number of sub-apertures.

In this work, a similar statistic is generated, based on the residuals between fiducials.

The stitching concept is now extended to have a wavefront sensor mounted on a real multi-axis translation stage – in the simplest case this may be a linear x-y stage that could be used for stitching measurements on flats only. The part to be measured is translated underneath the metrology instrument according to the layout of sub-apertures chosen. Figure 2.1 details such a scheme for testing large flats using a wave-front sensor with a collimated beam. Bray has previously demonstrated a stitching system for the measurement of flats based upon an instrument of this type and the mathematics of equation 2.3 on page 87, which Bray refers to as the tip-tilt-piston stitching function and claims a total repeatability of 2 nm RMS over an area of 100×100 mm on the system described in his paper [66].

2.1.2 Least-Squares Approach

A real metrology system based on the schematic in figure 2.1 will have many error sources such as noise, tip, tilt, piston, geometric distortion and polar rotation to name just a few. Furthermore, turbulence and vibrations will also introduce noise

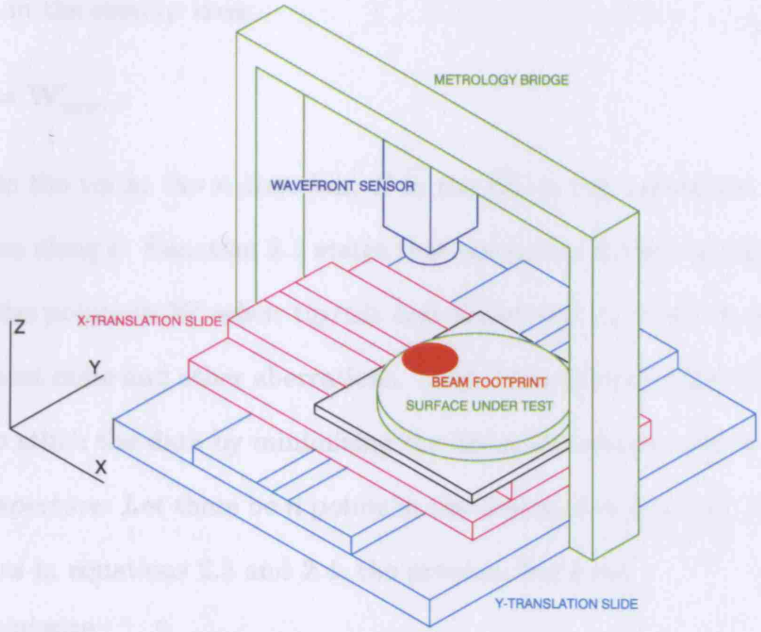


Figure 2.1: Details of a metrology stitching system for large flats. The x- and y-axes would normally be encoded for positional feedback to the stitching software. This is the essence of the metrology platform of Bray [66].

into the measurement result. The net result is that the stitching can never be exact.

A statistical approach can be used to find the best stitching transformations such as a least-squares method, and the determination of the transformations is called finding the *stitching parameters* for a given pair of sub-apertures.

Otsubo et al [1] described a process for correcting tip/tilt and piston terms between sub-apertures for the measurement of large flats using a least-squares approach similar to equation 2.5. The two sub-aperture measurements are labelled W and W' as before, and the physical displacement between them is $(x_0, y_0, 0)$. Furthermore, let the j$^{\text{th}}$ column in $\mathbf{W}'$ be labelled $\mathbf{W}'\{j\}$. The stitching of the points in the sub-apertures is effected by applying the following:

$$\mathbf{W}'_{\text{new}}\{j\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ A & B & 1 \end{bmatrix} \cdot \mathbf{W}'\{j\} + \begin{pmatrix} x_0 \\ y_0 \\ C \end{pmatrix} \tag{2.3}$$

such that in the overlap area,

$$\mathbf{W} = \mathbf{W}'_{\text{new}}, \tag{2.4}$$

where A is the tip in the x-direction, B is the tilt in the y-direction and C is the piston term along z. Equation 2.3 states that the points in the overlap area, $\mathbf{W}'$ are equal to the points in $\mathbf{W}$ when tip/tilt and piston are corrected in the absence of measurement noise and other aberrations. Least-squares can be used in the presence of noise to stitch the data by minimising the distances between equivalent points in each sub-aperture. Let there be n points in the intersection area and using the same variables as in equations 2.3 and 2.4, the problem becomes:

$$\underset{\text{over } \mathbf{W} \cap \mathbf{W}'}{\text{minimize}} \sum_{j=1}^{n} \left|\mathbf{W}\{j\} - \mathbf{W}'_{\text{new}}\{j\}\right|^2. \tag{2.5}$$

The 'best-fit' stitching between the two sub-apertures is performed by finding A, B, C, x_0 and y_0 in equation 2.5 to give the smallest residual. The idea can be extended to any number of sub-apertures and therefore allows, at least in theory, surfaces of infinite extent to be measured. The only issue here is how to define accurately the points in the overlap area.

Sub-aperture stitching is therefore an optimization problem and is often formed of many dimensions (or variables) not limited to A, B and C in equation 2.5.

2.1.3 Building a Stitching Algorithm

A simple stitching algorithm can be built up visually using a flowchart and the concepts covered so far. As the algorithm builds up a point will be reached where each operation will be represented in computer code and mathematics: a tangible product of this theory. This has the distinct advantage that the problem can be broken down into more manageable sub-problems each with well-defined inputs and outputs.

Figure 2.2 shows a generalized flowchart for the scheme described in this section

which contains high level tasks that will be broken down into discrete steps as the fledgling algorithm is developed in the coming sections.

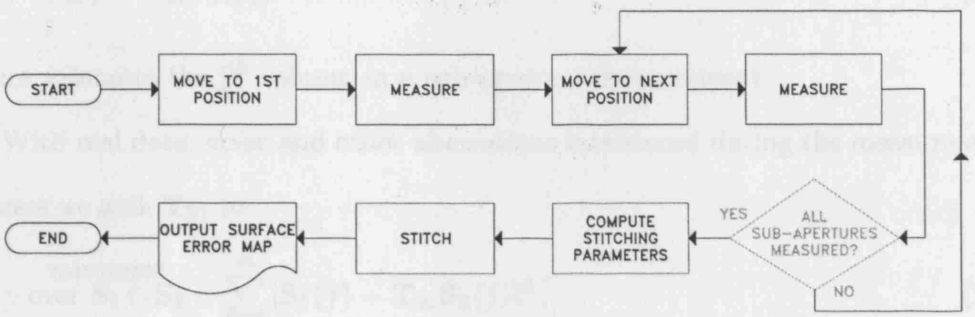


Figure 2.2: Flowchart for the first generation simple stitching algorithm.

In the measurement block of figure 2.2, a wavefront sensor will be moved relative to the stationery optic or vice-versa. By consideration of the sensor-part geometry a suitable arrangement of sub-apertures can be devised to measure the part in the optimal way. This sub-aperture arrangement has been called the *lattice map* by Murphy et al [71] in the *Spherical Stitching Interferometer* manufactured by QED Technologies, as described in figure 1.16 on page 77. Consideration of this *lattice map*, or sub-aperture map, allows a corresponding *fiducial map* to be produced detailing the positions of the physical fiducials. The requirements for the number and arrangement of fiducials is discussed in section 2.1.5

The method of stitching used in this work takes a slightly different approach to the methodologies described here and in chapter 1. The variable W is changed to $\mathbf{S}_i$ to represent the array of real surface data in the i^{th} sub-aperture measurement:

$$\mathbf{S}_i = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \\ z_1 & z_2 & \dots & z_m \end{bmatrix}, \qquad i = 1 \dots n, \tag{2.6}$$

where n is the number of sub-apertures and m is the number of points in the i^{th} sub-aperture. We consider two sub-apertures and let $\mathbf{T}_{21}$ be the $\mathbb{R}^3$ transformation matrix required to stitch sub-aperture $\mathbf{S}_2$ onto $\mathbf{S}_1$. With perfect data and a perfect

transformation, in the overlap area:

$$\mathbf{S}_1\{j\} = \mathbf{T}_{21}\mathbf{S}_2\{j\}, \tag{2.7}$$

where j denotes the j$^{\text{th}}$ column in a sub-aperture measurement.

With real data, noise and other aberrations introduced during the measurement process we seek $\mathbf{T}_{21}$ to

$$\underset{\text{over } \mathbf{S}_1 \cap \mathbf{S}_2}{\text{minimize}} \quad \sum_{j=1}^{m} |\mathbf{S}_1\{j\} - \mathbf{T}_{21}\mathbf{S}_2\{j\}|^2, \tag{2.8}$$

where the vertical bars indicate the l^2 norm defined by,

$$|\mathbf{x}| = \sqrt{\sum_{k=1}^{3} x_k^2}, \tag{2.9}$$

and $\mathbf{S}_1 \cap \mathbf{S}_2$ can be taken to mean the common points between the sub-apertures.

The exact nature of $\mathbf{T_{21}}$ and how it is computed will be defined in chapters 3 and 4 respectively. It is also noted here that a similar equation applies for for each set of overlapping sub-apertures, not just the first two and that the global sum of the residuals from each set of overlapping sub-apertures is minimized.

$$\underset{\text{over } \mathbf{S}_k \cap \mathbf{S}_l}{\text{minimize}} \quad \sum_{j=1}^{m} |\mathbf{S}_1\{j\} - \mathbf{T}_{21}\mathbf{S}_2\{j\}|^2 + |\mathbf{S}_1\{j\} - \mathbf{T}_{31}\mathbf{S}_3\{j\}|^2 + \ldots \tag{2.10}$$

The discussion used here is used to simplify writing down the mathematics where, otherwise, the indexing would become intrusive and unwieldy.

When m is large, there are many terms to consider for the least-squares optimization. Minimizing over a select number of points, *such as fiducial points only*, has the advantage that less computer time is required to solve the problem provided there are sufficient points and that they can be selected from the sub-aperture data. A further advantage is that it is not necessary to compute exactly which measurement points lie inside the overlap area.

2.1.4 Why Stitch?

Previously cited work on sub-aperture stitching has evolved from the need to overcome aperture size and also slope measurement limitations of traditional interferometry. Sub-aperture stitching therefore offers several benefits to the metrologist:

- The whole of a larger surface can be measured at a higher spatial resolution than would otherwise be possible. For example, if a beam expander was used to measure a large flat, then the lateral resolution obtained would be less that that obtained from a series of sub-aperture measurements without the expander.

- The size of the surface to be measured is limited only by the mechanics of translating the metrology instrument over the surface to be tested and the amount of memory needed to store the data arrays containing the sub-aperture data.

- Smaller diameter plano optical transmission elements can be used in the metrology instrument reducing capital equipment costs.

- It is possible to stitch complex aperture shapes by stitching where they could not be measured by a single full-aperture measurement. This is important in interferometry where the aperture is of an irregular shape that is incompatible with the phase unwrapping method.

- Small portions of a surface that as a whole departs significantly from a sphere may be tested using a sub-aperture measurement. The slope change over the whole surface may be too great to test full-aperture, but the slope change over small parts of the surface may be accommodated by sub-aperture interferometric testing.

There are of course several disadvantages

- Stitching inevitably introduces a further source of error to the measurement.
- Complicated processing is required to register the sub-aperture maps - the stitching algorithm.
- When the number of sub-apertures increases, the scope for stitching error also increases.
- Establishing the traceability of the result is more difficult using stitching.

2.1.5 Stitching with Fiducials

The decision to fiducialize the part was taken at an early stage in the OMAM project [76]. It was proposed that this would not only overcome the standard problem of establishing the part's orientation with respect to the metrology frame, but as we have seen could also be used to establish a set of known points in each overlap area.

We now consider how to arrange a set of fiducials in an overlap area between two sub-apertures. Once we have considered the theory, a method can be laid down to work with real instruments and real fiducials.

- A single fiducial mark placed in the overlap area of two sub-apertures will define a unique point where the sub-apertures stitch. However, the positions of the sub-apertures can still be arbitrarily oriented about this point.
- The application of two fiducials still does not solve the stitching problem since the sub-apertures can be tilted arbitrarily about the axis defined by the line joining the two fiducials. The sub-aperture positions are underconstrained and a further constraint is required to stitch them.
- The application of three **non-co-linear** fiducials fully constrains the positions of the two sub-apertures and allows errors such as tip/tilt and piston to be

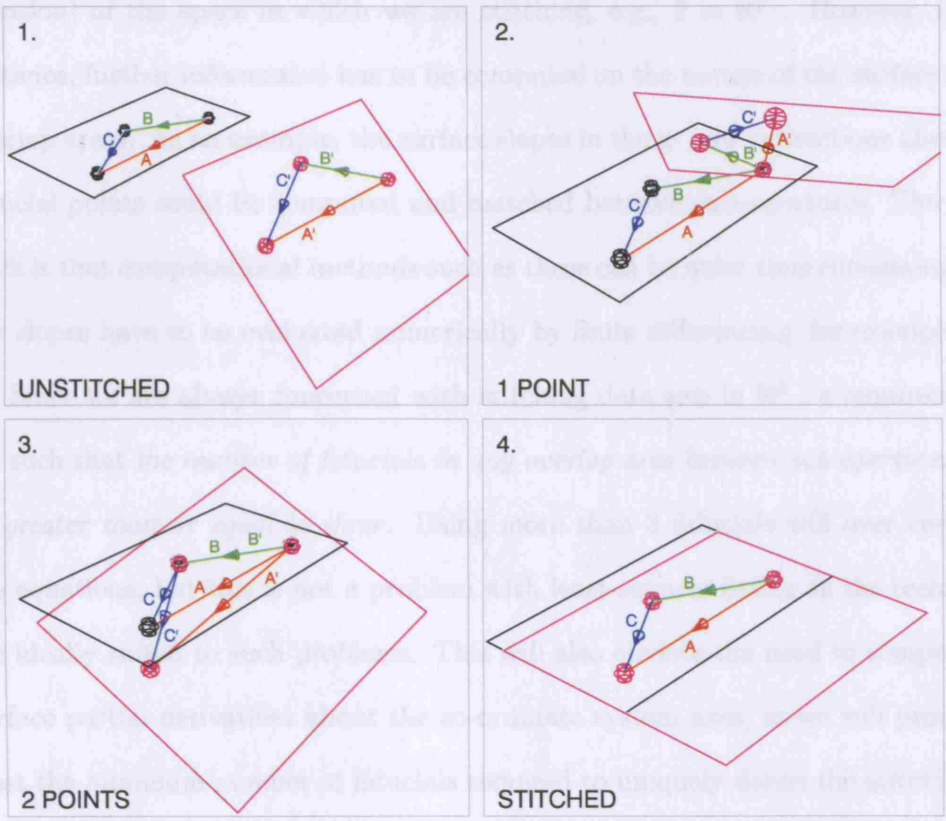


Figure 2.3: Steps showing the piecewise stitching of the magenta sub-aperture patch onto the black sub-aperture patch. The square outlines represent the data arrays of the two sub-apertures. (1): Unstitched sub-apertures with three fiducial points in the overlap area. (2): Stitched at a single point. (3): Two fiducials do not completely specify the stitching and a third is required. (4): Stitched sub-apertures showing the co-incidence of all three fiducial pairs.

eliminated. A good mathematical analogy of this is the use of three points to define a unique plane in $\mathbb{R}^3$.

- The application of more than three fiducials over constrains the sub-aperture positions but in the presence of noise allows for the possibility of smaller errors than with just three fiducials.

The analytical use of three fiducials is shown in figure 2.3 where the stitching of two sub-apertures is illustrated to completely specify the ideal stitching solution.

Stitching is also possible using any number of fiducials less than the order (di-

mension) of the space in which we are stitching, e.g., 2 in $\mathbb{R}^3$. However, in this instance, further information has to be computed on the nature of the surface in the overlap areas. As an example, the surface slopes in the x- and y-directions about the fiducial points could be computed and matched between sub-apertures. The drawback is that computational methods such as these can be quite time consuming since the slopes have to be evaluated numerically by finite differencing, for example [77].

Since we are always concerned with stitching data sets in $\mathbb{R}^3$, a requirement is set such that *the number of fiducials in any overlap area between sub-apertures must be greater than or equal to three.* Using more than 3 fiducials will over constrain the equations, but this is not a problem with least-squares fitting as the techniques are ideally suited to such problems. This will also obviate the need to compute the surface partial derivatives about the co-ordinate system axes, as we will provide at least the minimum number of fiducials required to uniquely define the stitching operations. With only two fiducials we can make two overlapping patches stitch about a line, but further information is required to correctly tilt them about the line to make them fully stitched.

Stitching can now be defined here as *transforming a set of surface data containing fiducial points, whose coordinates are known, onto another set of adjacent and overlapping surface data containing the same fiducial points, such that the sum of all the distances between equivalent fiducials is minimized.* This sum shall be termed the *fiducial residual.* This process shall be performed an all sub-apertures simultaneously to achieve the minimum sum of the distances between all fiducial pairs.

2.2 Metrology Devices and Fiducials

After establishing that at least three fiducials will be placed in the overlap region and that a least-squares minimization will be used, the physical process of metrology and the locating of the fiducial coordinates in the metrology device's output are now

considered.

Let the process of elucidating the coordinates of the fiducials be called *fiducial centroiding*. This process of fiducial centroiding may preclude the use of certain metrology devices and this section begins with some remarks on the constraints imposed by using fiducials.

- Placing physical marks on the surface under test implies that they must somehow be visible in the data from the optical metrology device.

- The fiducials must not affect the metrology of the surface measurement at any points other than at those corresponding to the points on the surface beneath the physical fiducial. The presence of a fiducial will result in a discontinuity in the phase and/or amplitude of the reflected wavefront. If small discontinuities are a problem for a given wavefront sensor then the use of surface fiducials may preclude the use of that wavefront sensor.

- The full centroid of every fiducial used must be available in the x-, y- and z-coordinates. If the z-coordinate is not directly available from the data some means of inferring it or computing a reliable estimate must be devised. This is relevant where something is attached to the SUT which modifies the phase data such that the actual height of the surface at the fiducial is not known.

- If the fiducials are not present in the output data of the metrology instrument then an image of the surface under test is required from the instrument with the same image-object mapping as the phase measurement result.

- A fiducial mark should be small in comparison to the size of the sub-aperture being measured.

The simplest method of fiducializing is to plot removable ink spots onto the surface under test. It was suggested by Peggs that in an automated CNC machine

tool system, an ink jet printer head could be used to apply well defined ink spots to the surface under test in pre-determined positions[1].

Henceforth it is to be taken that the fiducials will take the form of ink spots on the surface under test. The work of integrating surface fiducialization with a metrology platform is dealt with by Callender [79]. However, the work in this thesis is self contained in that an alternative manual fiducialization technique was ultimately used.

Fleig et al. have reported stitching sub-aperture data, using their Spherical Stitching Interferometer [loc cit.], without fiducials using a least-squares technique coupled to a precision 6-axis translation system with a very precise coordinate read-out system feeding into the stitching algorithm. The machine is calibrated before use and the calibration data is fed into the algorithm and used to obtain an initial transformation matrix for each sub-aperture before optimization.

The use of fiducials in this approach enables lower quality position encoders to be used on a translation, system since for flats, spheres and mild aspheres, it is not a requirement to know the position of the WFS to a very high degree of accuracy[2]: the fiducial points will define the relative positions of the sub-apertures provided that they can be centroided with sufficient accuracy.

The initial parameters for the stitching can still be estimated by consideration of the configuration of the translation system as measured by a series of position or angle encoders. These parameters can then be used as a seed for the transformation problem before numerical optimization takes place.

The previously mentioned OMAM PDWFS[3], dealt with in the next sub-section, has the advantage that it was designed to deal with discontinuities and scintillation in the wavefronts to be measured. The presence of fiducials on the surface definitely

[1]Intellectual property was sought for this method by NPL [78].

[2]This assumes a null test configuration.

[3]Chapter 1, page 69

causes discontinuities and possibly scintillation in the reflected wavefront and, in theory, this should not have posed any problem for this type of wavefront sensor.

2.2.1 OMAM: From Curvature Sensing to Phase Diversity

Curvature

Wavefront curvature sensing was first used by Roddier et al. to test mirrors after their installation into telescopes [50, 51].

In Roddier's approach, the wavefront intensity was sampled in two planes spaced symmetrically about the telescope's entrance pupil as explained in section 1.6.2. An estimate of the axial intensity gradient was formed by a point by point difference between these images which Roddier showed could be related to the second derivative of the wavefront phase. Thus by solving a Poisson equation similar to equation 1.12, the phase of the incoming wavefront could be established and used to check the wavefront reflected from the telescope primary. Roddier's original approach to curvature sensing used multiple imaging devices to obtain the two intensity images and a schematic of this scheme is shown in figure 2.4.

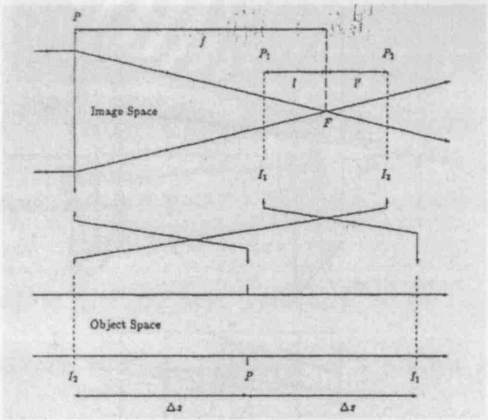


Figure 2.4: Taken from [80] showing the arrangement used in classical curvature sensing. "In the image space (top), the recorded illuminations I_1 and I_2 appear as defocused stellar images. In the conjugate object space (bottom), they appear as defocused pupil images."

Woods and Greenaway extended the idea and presented a method that gave the

form of the incoming wavefront directly in terms of the Zernike polynomials via a matrix inversion [55].

More recent work by Blanchard and Greenaway involves using a type of zone plate to provide multi-plane imaging onto a single CCD camera [52] to avoid the need for two imaging devices. They fabricated binary phase diffraction gratings in the form of off-axis zone plates to apply defocus to the incoming wavefront. One order of the grating applies positive defocus, whilst the opposite order applies negative defocus. In this way, two planes separated along the optical axis can be imaged onto a single image plane. Figure 2.5 shows a grating pattern used to provide defocus from the literature and figure 2.6 outlines the principle of multiplane imaging.

Figure 2.5: Typical defocus grating from [52]. The grating is essentially an off-axis zone plate.

Combining a diffraction grating and a software engine implementing the theory of curvature sensing allows the construction of a simple and effective wavefront sensor with high acquisition rates, particularly if only low-order terms of the Zernike expansion of the incoming wavefront are required. However, the mathematics used assume that the wavefront is continuous and is circular - a condition of using Zernike polynomials. If this is not the case then the output data are not valid [81].

The defocus grating is the subject of a patent and a commercial device using the

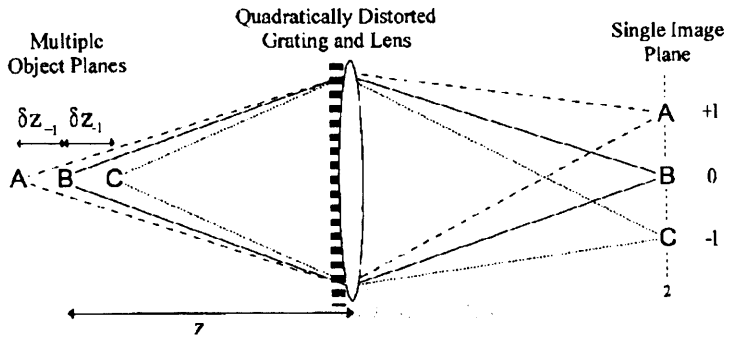


Figure 2.6: Principle of the imaging produced by the grating shown in figure 2.5. The multiple object planes are imaged into the grating -1, 0 and +1 orders on a single plane.

grating is currently marketed by Arden Photonics Ltd, the Arden AWS-50 wavefront sensor.

Phase Diversity

The Heriot-Watt led OMAM PDWFS was to extend upon the curvature system and enable a more general approach to wavefront sensing termed Phase Diversity WFS. Key features are arbitrarily chosen aberrations that can be encoded onto the incoming wavefront, and a new phase recovery algorithm. The new algorithm was designed to overcome some of the limitations of the previous curvature algorithm of Woods and Greenaway [loc cit.], namely the use of Zernike polynomials, and to provide a direct computation of the wavefront phase.

The OMAM PDWFS uses a binary phase diffraction grating, as with the basic curvature WFS, that introduces defocus and one other general precomputed aberration onto the incoming wavefront - hence the term phase-diversity. For example, experiments at Heriot-Watt have used gratings that introduce defocus, astigmatism, coma, trefoil and primary spherical aberration [54]. The aberrations are encoded onto the grating orders – the zeroth order being the original unaberrated wavefront, the positive orders encoding $n \times m$ waves of the aberration onto the wavefront and the negative orders encoding $-n \times m$ waves of the aberration, where n is the grating

order and m is another constant specifying the strength of the aberration.

Heriot-Watt were to deliver a working prototype PDWFS to UCL in May 2004. After some initial schedule slippage the author and a colleague worked at Heriot-Watt (H-W) to build a prototype WFS to allow an assessment to be made of the size and weight of such a device whilst the difficult mathematical aspects were addressed by H-W. The details of the UCL-built prototype PDWFS can be found in appendix C.

According to Heriot-Watt the PDWFS can be tailored to the metrology requirements by varying the properties of the diffraction grating to make it more sensitive to particular aberrations [82]. This fact was to make the PDWFS a key technology in the whole OMAM project. However, H-W encountered unexpected problems with the mathematics of phase retrieval and were not able to deliver a working system to UCL during the project. After slipping to May 2004, the delivery milestone slipped to July 2004 and then further, and at this epoch the non-delivery became a serious impediment to this work and an alternative wavefront sensor was sought. Lack of funding meant that the only viable options were existing optical metrology devices owned by UCL and by NPL.

The work presented here represents a work-around for the non-delivery of the PDWFS, embodying all aspects of surface fiducialization and geometric transformations envisaged for the PDWFS.

During the project a Fisba μPhase interferometer was loaned to UCL by NPL under an agreement for two years and was identified as a possible replacement for the PDWFS. Furthermore, UCL's Wyko 6000 laser interferometer was also identified as a possible candidate for trials with metrology data stitching. However, it is unreasonable to consider just the Wyko and Fisba systems as possible candidates without considering other alternatives. Therefore the following section provides a comparison of some of the more common metrology devices considered for the work-

around.

2.3 Selection of an Alternative WFS

A number of technologies were available to the project including the defocus grating curvature sensor from Heriot-Watt and the existing metrology instruments at UCL and NPL. A traditional compliance matrix approach was taken to evaluate the merits of each instrument in turn. The matrix has binary elements (as yes/no) and the rows are summed to give the overall compliance – the instrument with the highest compliance is the most appropriate. The matrix is shown in table 2.1 and explained as follows. The first column simply names the metrology instrument whilst the

Device	Fid Centroid Data	Config Optics	Light-weight	Fids in Data	Devel	Own/ Acquire	Score
Compact phase-shifting interferometer	Yes	Yes	Yes	Yes	Yes	Yes	6
Phase-shifting interferometer	Yes	Yes	No	Yes	Yes	Yes	5
H-W curvature WFS	No	Yes	Yes	No	Yes	Yes	4
OMAM PDWFS	Possible	Yes	Yes	Possible	Yes	No	4
Instantaneous Phase-shifting interferometer	Yes	Yes	No	Yes	No	No	3
Curvature WFS	No	Yes	Yes	No	No	Yes	3
Shack-Hartmann type	No	Yes	Yes	No	No	No	2

Table 2.1: Compliance matrix for several commercial and research wave-front sensing technologies considered in this work. The ranking proceeds from the top down.

second column indicates whether or not fiducial centroid data can be easily obtained by analysis of the instrument's output. Column three simply indicates whether a device can be easily configured with different transmission/focusing elements. An instrument is defined as lightweight in column four provided that its weight is less than 5 kg[4] and column five indicates if the co-ordinates of the fiducials can be

[4]From a constraint imposed by the swing-arm profilometer.

directly obtained from the instrument's output. The sixth column indicates if a software interface to the output data is already available and the seventh column indicates if UCL/OMAM already own the device or if it could be acquired for the project. Column eight gives the compliance as a sum of the rows (yes = 1, no = 0, possible = 0.5).

Figure 2.7 depicts, for comparison, some sample schematic diagrams of the main components of the more frequently encountered types of wavefront sensors considered here and referred to in the following paragraphs.

Compact Phase-Shifting Interferometer

At present, Fisba Optik manufacture a range of compact Twyman-green interferometers under the tradename μPhase. Their size and functionality makes them ideal candidates for this work. A description of the μPhase interferometer on loan to UCL from NPL is provided in section 2.3.2. Along with the other types of interferometer mentioned below, any marks or spots on the surface under test result in a corresponding identifiable region in the output data.

Phase-Shifting Interferometer

Commercial phase-shifting Fizeau and Twyman-green interferometers have been used in optical testing for many years and are well proven. The 6" Wyko 6000 interferometer of UCL meets all the criteria except for size and weight (57 kg). However, the software interface and quality of the instrument make it an ideal candidate for use in the development of the stitching algorithms. The instrument is described in subsection 2.3.3.

Heriot-Watt Defocus Curvature Sensor

A defocus grating and some analysis software was supplied by Heriot-Watt university for evaluation. They cannot be used for this project since discontinuities in the

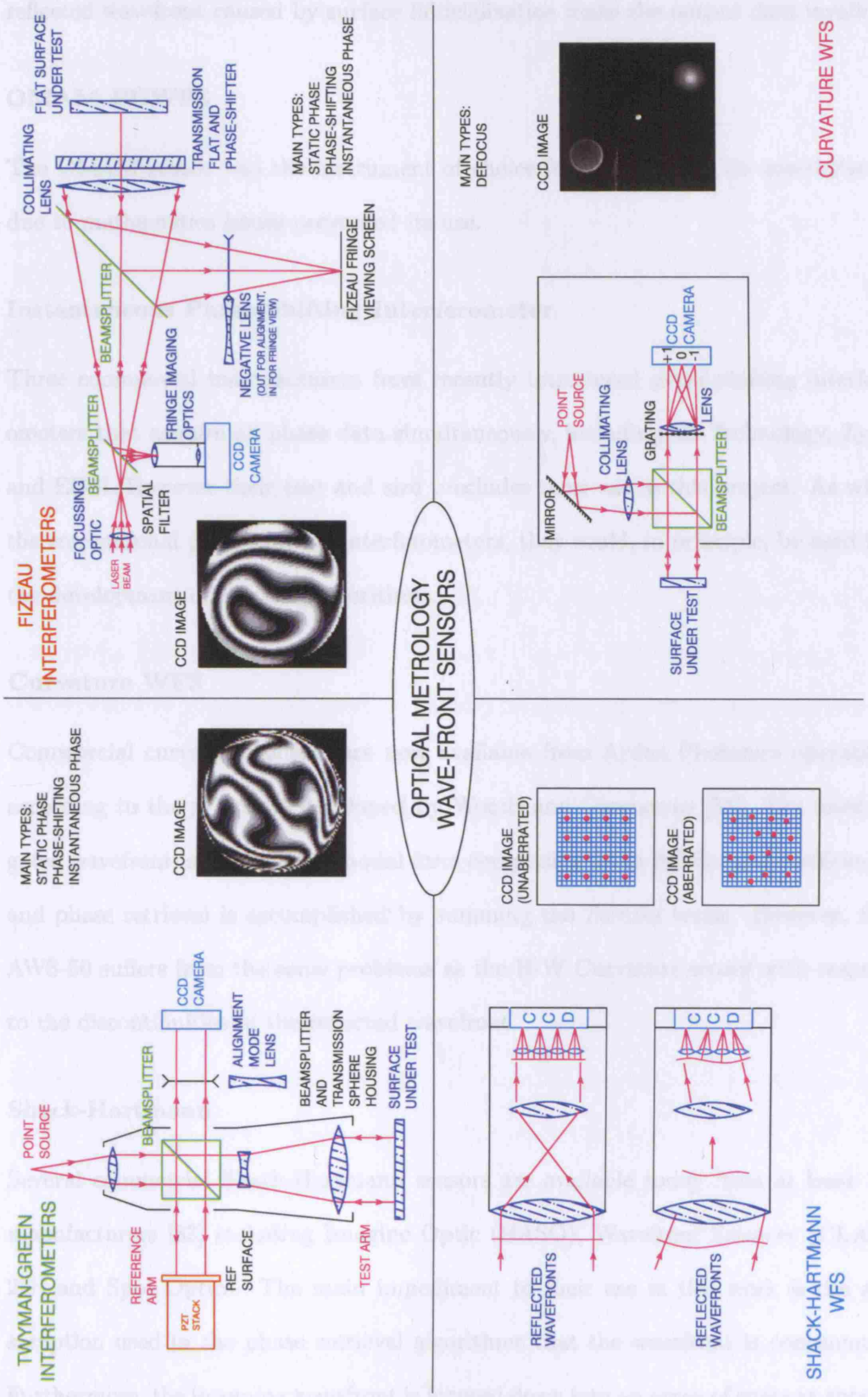


Figure 2.7: Common internal configurations and outputs for some common optical metrology tools.

reflected wavefront caused by surface fiducialization make the output data invalid.

OMAM PDWFS

The OMAM sensor was the instrument of choice for the project. Its non-delivery due to mathematics issues prevented its use.

Instantaneous Phase-Shifting Interferometer

Three commercial manufacturers have recently introduced phase-shifting interferometers that acquire all phase data simultaneously, including 4D-Technology, Zygo and EDSI. However their cost and size precludes their use in this project. As with the conventional phase-shifting interferometers, they could, in principle, be used for the development of stitching algorithms.

Curvature WFS

Commercial curvature sensors are now available from Arden Photonics operating according to the principles developed by Woods and Greenaway [55]. The interrogated wavefront is presented in modal form comprising up to 200 Zernike coefficients and phase retrieval is accomplished by summing the Zernike terms. However, the AWS-50 suffers from the same problems as the H-W Curvature sensor with respect to the discontinuities in the reflected wavefront.

Shack-Hartmann

Several commercial Shack-Hartmann sensors are available today from at least 10 manufacturers [83] including Imagine Optic (HASO), Wavefront Sciences (CLAS-2D) and Spot Optics. The main impediment to their use in this work is the assumption used in the phase retrieval algorithms that the wavefront is continuous. Furthermore, the incoming wavefront is focused down into an array of spots as shown in figure 2.7 on page 103. This means that no fiducial information is available from

a Shack-Hartmann WFS. Combined with cost, this fact rules out the use of such a device in this work.

2.3.1 Choice

With ready availability and the high scores in the compliance matrix of the Fisba μPhase and Wyko 6000 interferometers, it was decided that both interferometers should be used to develop the stitching software in the absence of the OMAM PDWFS. Both instruments output the same type of data, albeit in a slightly different format, that can be imported into user-designed and programmed software via a simple interface application.

The interferometers are able to deal with small discontinuities in the incoming wavefront and provide an established method for representing data points that are not part of the aperture or where no phase could be computed due to aperture discontinuities. This is dealt with more fully in section 2.5 where the fiducial centroiding algorithms are developed. The Wyko 6000 can be used to provide metrology for stitching by moving the part relative to the stationary interferometer if a suitable arrangement can be devised to translate the part. Section 6.4 details such a method used to measure a real part.

2.3.2 Fisba μPhase Interferometer

In late 2005 a mark I Fisba μPhase Twyman-Green phase-shifting interferometer was loaned to UCL by NPL for evaluation and as a possible substitute for the OMAM PDWFS. The instrument is extremely compact and occupies only $60\times30\times25$ mm^3 of space without reference optics and is thus ideal for use in on-machine metrology applications. The interferometer body consists of the mechanics required to contain the fibre-fed point source, the spherical reference surface, PZT stack and the CCD camera. A removable lens between the beamsplitter and CCD camera serves to focus the test-arm image to a point to enable the SUT to be correctly aligned with

the reference beam.

The optical source consists of a frequency stabilized HeNe laser operating at 632.8 nm feeding into a single mode optical fibre. The general layout of the interferometer is shown in figure 2.8. The Fisba system provides a maximum of 100x100 data points in the output phasemap. The interferometer was supplied to UCL with a test stand

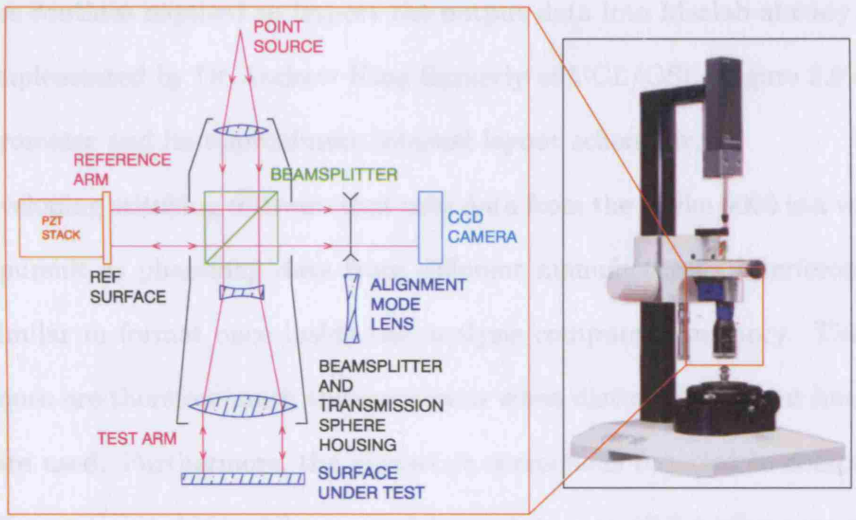


Figure 2.8: Fisba μPhase interferometer: schematic and image showing the reference and test arms and the phase-shifting PZT stack.

and a 10/∞ beamsplitter/collimator supplying a 10 mm $\varnothing$ collimated beam at the output. Two converging lenses were supplied that mate to the 10/∞ lens with 60° and 90° cone angles for the testing of spheres. Calibration of the interferometer was carried out using a 4" $\varnothing$, λ/20 P-V reference flat for the collimated beam, and a 10 mm radius sphere for the convergent beams. Custom software for importing the Fisba-derived data into Matlab was programmed by the author as listed in appendix E.

2.3.3 Wyko 6000 Interferometer

The Wyko 6000 Fizeau interferometer has a 6" aperture and a frequency stabilised He-Ne laser source operating at 632.8 nm. Reference calibration is provided by a 6"

$\varnothing$ reference flat to $\lambda/20$ P-V or to $\lambda/10$ P-V with the appropriate transmission and return reference spheres. UCL has a range of transmission spheres available for use including $f/11$, $f/3.3$ and $f/0.9$ and also several return spheres. Phase stepping is accomplished by translating the computer-controlled reference and requires 167 ms for a complete acquisition. One key advantage of using this instrument is that the software routines required to import the output data into Matlab already exist and were implemented by Dr Andrew King formerly of UCL/OSL. Figure 2.9 shows the interferometer and its approximate internal layout schematic.

Developing stitching software that uses data from the Wyko 6000 is a very worthwhile pursuit as phasemap data from different manufacturer's interferometers are very similar in format once inside the analysis computer's memory. The software techniques are therefore much the same even when distinctly different hardware devices are used. Furthermore, the geometric corrections required to compensate for the different transmission spheres used are common to all interferometers in terms of setting the base radius of curvature for the particular transmission sphere used. The Wyko 6000 is an extremely well known and respected instrument and its large aperture allows, for example, the efficient testing of optical flats in the 0.25 m^2 range. To illustrate this, consider the transmission flats available to the author for the Fisba and Wyko systems. The Fisba has a clear aperture of 10 mm whereas the Wyko has a clear aperture of 152 mm. To stitch a flat of 0.25 m^2 area will require approximately 5000 sub-aperture measurements with the Fisba without a beam expander, and just 25 with the Wyko. This argument is of course invalid when testing concave spheres, but a similar argument applies when testing convex spheres.

In contrast, the Fisba's extremely light weight and small physical dimensions make it ideal for mounting on to the end of the NPL/UCL swing-arm profilometer; with the Wyko at 57 kg, this is simply not possible. Callender [79] addresses the mounting of a compact interferometer onto the end of the NPL/UCL swing-arm.

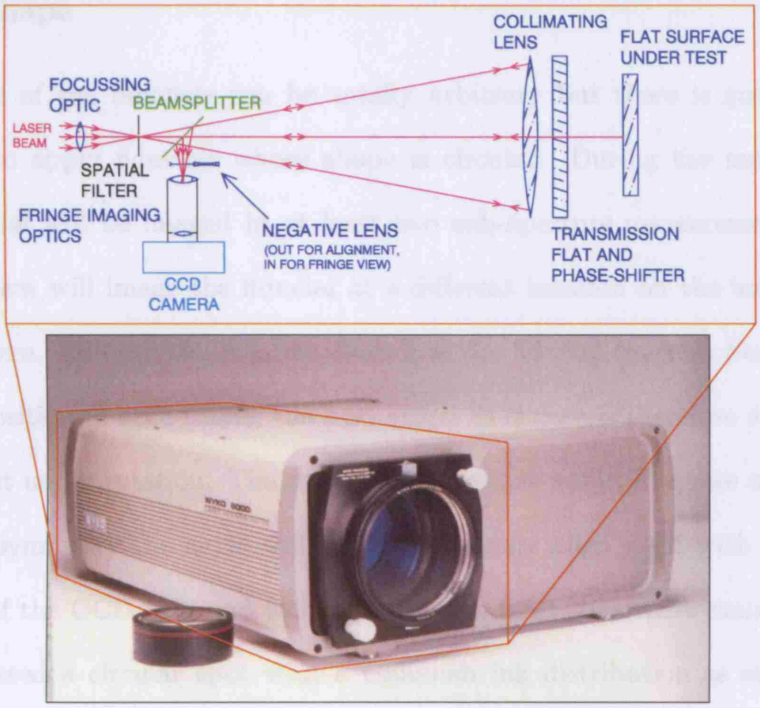


Figure 2.9: Wyko 6000 and schematic diagram. Coherent light from the point source is spatially filtered at the back conjugate of the collimating lens after passing through a beamsplitter. Collimated light emerges from the lens and is partially reflected by the transmission element. The transmitted beam is reflected from the surface under test back into the interferometer where it interferes with the beam reflected from the transmission element. The optics then focus the interference pattern onto the CCD camera for acquisition.

2.4 Surface Fiducialization

In this work, when stitching measurements of real parts, and after each sub-aperture map was designed, *fiducials were simply drawn on to the surfaces under test by hand using felt-tipped marker pens* according to the principles of subsection 2.1.5[5].

2.4.1 Shape

The shape of the fiducials can be totally arbitrary but there is good reason for choosing to apply fiducials whose shape is circular. During the metrology stage each fiducial will be imaged in at least two sub-aperture measurements and each measurement will image the fiducial at a different location on the interferometer's CCD camera. To have the greatest chance of the fiducial centroid being located in its true position in each frame, the best shape to choose is therefore a circle as this is invariant under rotation. The rationale being that unlike a square or other shape with less symmetry the circle will not preferentially align itself with the rows and columns of the CCD chip and will thus give the most consistent centroids. Walker has suggested a circular spot with a Gaussian ink distribution as such a fiducial may be centroided in an image of the surface under test using standard astrometry software [84]. This approach is not possible with the Wyko and Fisba interferometers as a pure image of the surface under test cannot be acquired or generated.

2.4.2 Size

The sizes of the fiducials can be constrained by consideration the interferometer's CCD resolution, the transmission sphere used, the radius of curvature of the part to be measured and the phase unwrapping method. This can only happen at the time when the full details of the test configuration are known. For example, when testing

[5]Callender [79] deals with an automated method of surface fiducialization and ink choice for a parallel stream of work under (C) in section 1.8.

a flat, the interferometer pixel size multiplied by the system magnification will give the linear resolution unit on the surface under test.

The situation for spheres is complicated by the f/# of the transmission element and the surface under test. Once the f/# are known the resolution can be calculated. As a rule-of-thumb, a fiducial should occupy at least 3×3 pixels on the interferometer CCD frame for centroiding purposes as anything smaller cannot be centroided reliably.

The maximum size of a fiducial is related to how the phase is computed by the interferometer and this depends on the particular model used. It has been found by experiment that provided the fiducial does not cause the reflecting aperture to be split into two then the phase can be computed for all other points on the surface for the Wyko and Fisba systems. Fiducials 12×12 pixels square and larger have been successfully imaged without causing data loss in other areas of the recovered phasemap.

2.5 Centroiding Fiducials in Interferometer Data

When a test wavefront is incident upon an optically reflecting fiducialized surface in an interferometric test, the following are asserted:

Thickness Part of the incident wavefront may be transmitted by an ink spot fiducial if it is optically transparent enough. The phase of the light as it leaves the spot after reflection by the surface will be determined by the distribution of thickness in the ink spot and its refractive index. Discontinuities are therefore expected in the returned wavefront phase in the regions corresponding to the positions of fiducials.

Absorption Part of the incident wavefront may be totally or partially absorbed by the ink pigment in the fiducial. We can therefore expect discontinuities in

the amplitude of the wavefront reflected from the optical surface in the regions corresponding to the positions of fiducials.

Scattering Part of the incident wavefront may be scattered by the ink in the fiducials causing a reduction in intensity in the returned wavefront at fiducial locations. This reduction in intensity may be below the threshold required for successfully interpreting the phase of the returned wavefront at the points in question.

From the above assertions it is argued that the most appropriate ink to choose for the fiducials is very absorbent of the 633 nm illumination used and also consists of very fine particles to scatter non-absorbed light in all directions. Furthermore, the results of some experiments to measure the thicknesses of some representative ink spots are shown in figure 2.10. A very quick qualitative analysis of the spots reveals that they

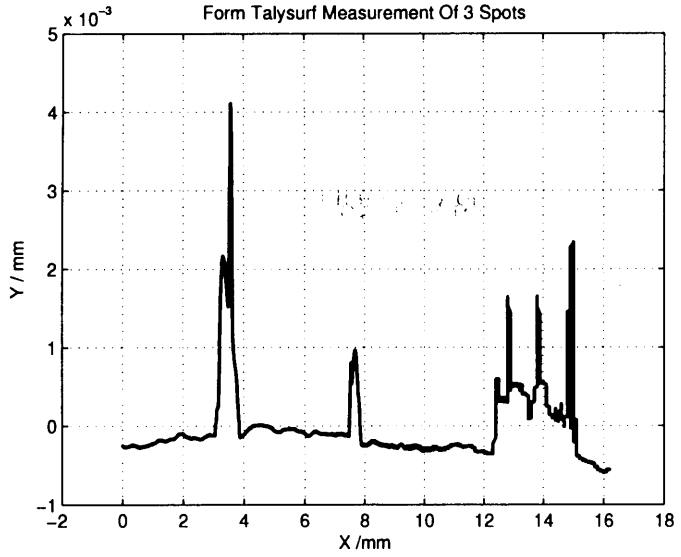


Figure 2.10: Form-Talysurf measurement of 3 ink spots from a pen used in the experiments.

have heights in the region of 0.5 – 2 μm which corresponds to approximately 1–4 waves at 632.8 nm. The spots also have sharp edges which suggests that a suitable discontinuity will be generated in the phasemap of the interferometer. The sharp

peaks in the profiles are the result of high ink concentrations in certain regions of the spots. Note that the centre spot has a width of significantly less than 1 mm which may be smaller than the resolution of the interferometer in some circumstances.

The following further assertions are made:

Placement The ink spots will be placed in precomputed positions. The positions may be simply derived by hand or by a software algorithm from a parallel stream of work.

Solvent The ink spots must be completely be removable and once a fiducial has been removed the surface should show no signs of it having been there at all. Pigment ink felt-tipped pens have been found to produce acceptable results as the solvent base is generally an alcohol and the pigments do not etch the surface.

It was found that translucent inks were less reliable because the optical phase thickness was not constant which leads to undesirable peaks in the data at the fiducial edges. Walker [85] first suggested the use of a UV fluorescent ink, but there are two reasons why this is not useful. Firstly, the inks usually emit in a given waveband and the radiation is not coherent and thus has no effect inside the interferometer except to increase noise. Secondly, the inks are also translucent and thus suffer from the same phase thickness problems as non-fluorescent translucent inks. The effects of the UV fluorescent inks can be seen in figure 2.11. To the left of the figure is the raw CCD data showing the interferogram. The variable optical thickness of the spots can clearly be seen. The resulting phasemap is shown to the right of figure 2.11 where the ink results in high spots around the edges of the fiducial. The nature of the data in the centre of the fiducial is discussed in the next section.

A black OHP pen was selected from the experiments since it has a matt black powdery ink pigment that was found to provide the most consistent results when

Figure 2.11: Left: Interferogram from the Fisba Twyman-Green interferometer, showing the presence of 3 fiducials made from UV fluorescent ink. Note the varying phase thickness of the ink. Right: Phasemap for the interferogram showing the regions corresponding to the fiducials where the phase could not be evaluated. In both images the banding is a diffraction effect caused by dirt in the interferometer optics.

taking repeat measurements of a fiducialized part. Figure 2.12 shows the results when using the OHP pen and the Wyko 6000 interferometer.

2.5.1 Practicalities

Fiducials were applied to a 4" diameter $\lambda/4$ RMS flat made from BK7 glass using the commercial OHP marker pen[6]. The effects of applied fiducials on the interferometer data can be clearly seen in figure 2.12. The fiducials are characterised in this data by their numerical value – the interferometers used assign invalid data points the value of the largest negative integer that can be represented in the computers word size to indicate '`NOT APERTURE`' points. These are points within the returned phasemap where no phase could be computed by the interferometer's data reduction software.

The fiducial pixels in the right-hand image of figure 2.12 have been assigned the `NOT APERTURE` value by the software since the phase could not be determined at those points. As the interferometer moves the reference surface in increments of $\pi/2$ between measurements, the fringes move and therefore the intensity mea-

[6]OHP Pen from Niceday by Guilbert. Type no. MR38E Black, non-permanent medium.

Figure 2.12: Left: Interferogram from the Wyko 6000 showing interference fringes and the presence of fiducial marks drawn on the surface (black spots). Right: Phasemap corresponding to the interferogram left, showing clean holes in the data at the fiducial locations. The exact nature of the data in the fiducial regions in the phasemap is described in the text and has an effect on how the fiducial centroids are computed.

sured by any pixel on the CCD camera changes. If the change in intensity is below a preset value then the interferometer software assigns the resulting data point in the phasemap the value of `NOT APERTURE`. This indicates to the software that that particular value is not to be used in the computation of the surface statistics and is therefore a convenient way to determine where the edges of the surface measurement lie. Otherwise, points inside the aperture that are assigned this value must correspond to locations on the surface under test from which no wavefront was returned, such as the fiducial points.

The phasemap array data is then read into Matlab by a routine provided by Dr Andrew King, formerly of UCL/OSL, where each `NOT APERTURE` value is replaced by the `NaN` constant. The `NaN` constant is used in information science to denote values that are not to be treated as being numerical. This provides a standard way in higher level languages of distinguishing between valid and invalid data points. Furthermore, Matlab, the software of choice for this work provides a very fast and convenient function for dealing with `NaN` entries in sets of data. Since each recorded phasemap will contain many instances of the `NaN` constant this is a distinct advantage

where speed is concerned.

The general form of a phasemap once imported into Matlab is a rectangular array of numbers interspersed with NaN constants for pixels outside the aperture, or pixels where no phase information could be derived such as regions corresponding to where the fiducials were present. Table 2.2 shows a mock phasemap 10×10 pixels in size with a circular aperture and a fiducial located in the centre consisting of 5 pixels. The plot is shown in figure 2.13, but note that piston has not been removed (i.e., the average of this mock data is not zero). The important features present in phasemap

NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
NaN	NaN	NaN	3	4	2	2	NaN	NaN	NaN
NaN	NaN	3	3	4	5	4	4	NaN	NaN
NaN	2	2	2	4	4	3	3	3	NaN
NaN	3	3	4	NaN	NaN	6	5	4	NaN
NaN	2	2	2	NaN	NaN	NaN	6	4	NaN
NaN	2	3	4	3	4	5	5	5	NaN
NaN	NaN	3	3	4	5	4	4	NaN	NaN
NaN	NaN	NaN	2	2	1	2	NaN	NaN	NaN
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Table 2.2: Mock phasemap containing a fiducial. This data contains a DC (piston) term since the mean value of the phase is 3.

data are named in figure 2.14 to be considered in anticipation of developing a fully automated acquisition and centroiding algorithm. The key features are:

Not Aperture This is the region of the phasemap outside of the area where interference was occurring. The NaN constant is used for data in this area.

Aperture Data This is the region where interference was occurring and a phase for the interference fringes was calculated at each point. The aperture data

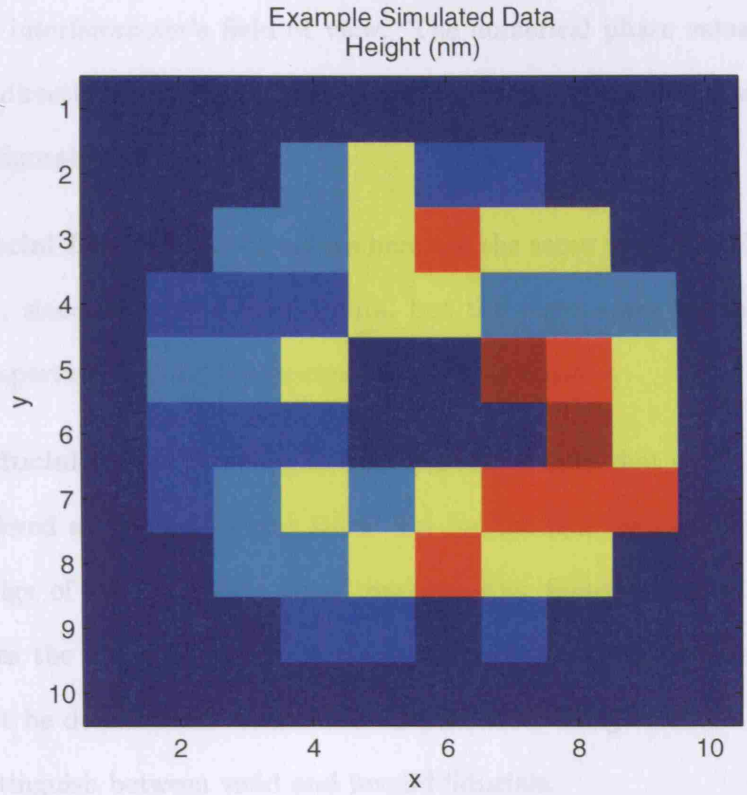


Figure 2.13: Plot of the mock data from table 2.2

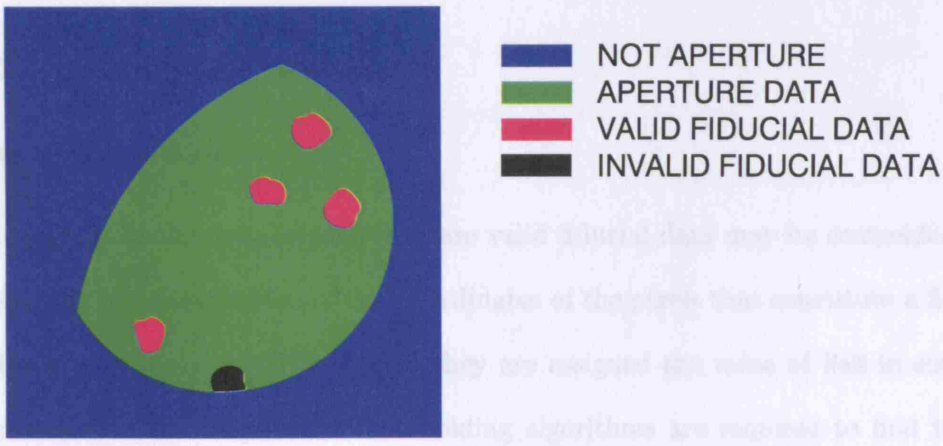


Figure 2.14: Terminology for features in a rectangular array of phase data from an interferometer. Note the size of the fiducials has been exaggerated.

area may have an irregular shape if the surface under test overlaps the edge of the interferometer's field of view. The numerical phase value is generally given directly in nm for the interferometers used for this work, although this is configurable for the Fisba.

Valid Fiducial Data The data values here are the same as in the 'Not Aperture' region, since no interference occurs, but the regions are located within the main aperture and are hence considered as fiducials.

Invalid Fiducial Data These regions correspond to data that would be otherwise considered as 'Valid Fiducial Data' but for the fact that they intersect with the edge of the 'Aperture Data' region. Any fiducial located such that it touches the outer boundary of the aperture must be ignored as its full size cannot be determined. Automatic fiducial centroiding software must be able to distinguish between valid and invalid fiducials.

The regions containing `NaN` inside the aperture can be detected by software and centroided using the method described in sub-section 2.5.2. This feature is common to both the Wyko and Fisba interferometers used in this work.

2.5.2 Centroiding Method

The x- and y-coordinates

The regions inside the aperture that are valid fiducial data may be centroided geometrically by consideration of the coordinates of the pixels that constitute a fiducial - this is made easy by the fact that they are assigned the value of `NaN` in software and therefore no complicated thresholding algorithms are required to find the regions. Each fiducial pixel will therefore be weighted equally in the evaluation of the centroid such that the x and y coordinates of the centroid are just the averages of the pixel coordinates of the fiducial in the two axes. Let x_f and y_f be the pixel co-

ordinates in the ith fiducial containing n pixels. The fiducial centroid is thus defined as:

$$C_i = \left(\frac{\sum_{f=1}^{n} x_f}{n}, \frac{\sum_{f=1}^{n} y_f}{n}, z \right), \tag{2.11}$$

The manual fiducial centroiding software developed in this work as part of the stitching algorithm allows the user to point to each fiducial in turn on a plot of the phasemap to assign a number to each one before the centroid is computed. An automatic search is then made in a circular region around the points picked to determine the exact fiducial centroid by averaging the coordinates of the `NaN` entries. This procedure is acceptable if a only a small number of fiducials is used – a fully automated system is demonstrated in section 2.7 that can be used in the presence of a large number of fiducials.

The z-coordinate

The fiducial z coordinate is found by consideration of the data around the fiducial. Here, the symmetry of a circle is used again, or more specifically, that of an annulus. The software algorithm developed places an annulus at each fiducial's centroid and the real data points lying underneath the region of the annulus are averaged in z (phase) to give a value for the z-coordinate of the fiducial as $z = \frac{\sum_{\text{annulus}} \phi_i}{n}$, where n is the number of points in the annulus. The minor radius of the annulus is chosen to be just larger than the radius of the fiducials and the major radius is chosen to be between 3 and 10 pixels larger than the minor radius. The value for z then occurs at the annulus centre – precisely the x-y centroid of the fiducial. Figure 2.15 illustrates the process of finding z.

Once all fiducials in a sub-aperture measurement have been found, their centroids are appended to the sub-aperture data array shown in equation 2.6 to give the full

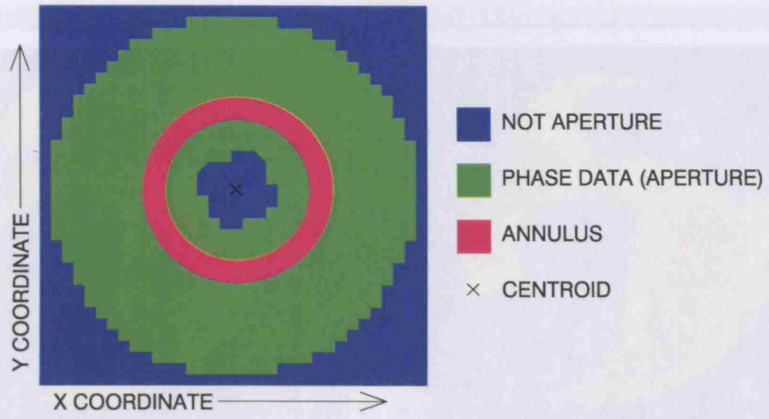


Figure 2.15: Blue: not aperture. Green: phase data. Magenta: annulus of points to average to elucidate z.

sub-aperture array (with a slight change of index variables):

$$\mathbf{S}_i = \begin{bmatrix} x_1 & x_2 & \dots & x_m & fx_1 & \dots & fx_n \\ y_1 & y_2 & \dots & y_m & fy_1 & \dots & fy_n \\ z_1 & z_2 & \dots & z_m & fz_1 & \dots & fz_n \end{bmatrix}, \tag{2.12}$$

where the fx, fy, and fz are the fiducial x, y, and z coordinates respectively.

Since an annulus is invariant under rotation about an axis perpendicular to its area, the results will be consistent between sub-apertures and in theory one could fit any form to the points of the annulus and achieve consistent results between sub-apertures - taking the mean phase is just the simplest method to achieve a consistent result. A portion of the Matlab manual centroiding software window is shown in figure 2.16 that illustrates the operation.

Matching Sets of Points

The tip/tilt and piston terms between two overlapping sub-aperture measurements will never be zero. Therefore we cannot expect the z-coordinate of a given fiducial to be identical in two (or more) sub-apertures. What is important is *the coordinates of the fiducials relative to each other in a given sub-aperture, and the tip/tilt between the sub-apertures*. An analysis of this is made in section 2.6 to assess the probable

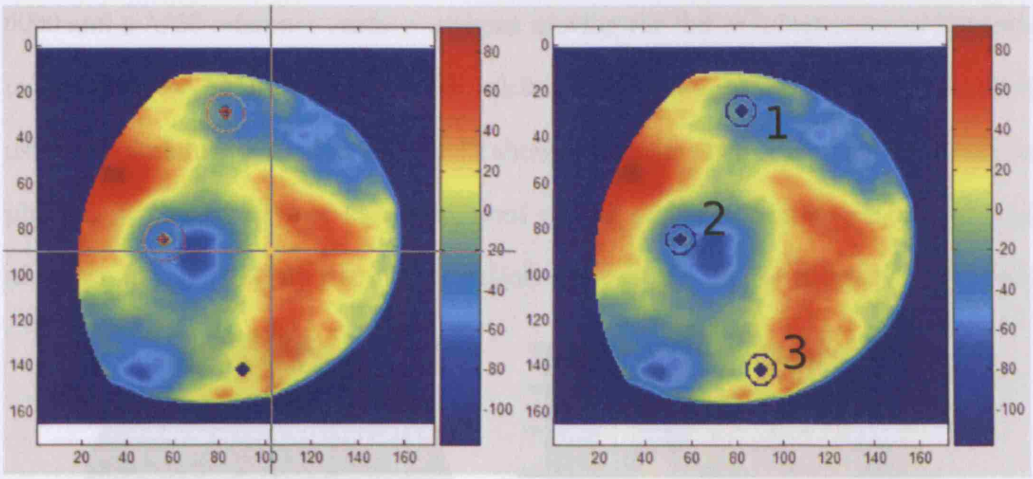


Figure 2.16: Matlab manual centroiding software window portion: Wyko interferometer data showing regions of `NaN` inside the main aperture corresponding to the fiducial positions. Left shows the user cursor used to specify a circular region in which a fiducial lies (red circles). Right shows data after centroiding using the annulus method. Blue coloured points are the ones used to calculate z.

errors to be expected in stitching.

The software code developed for centroiding fiducials in Matlab array formatted interferometer data is outlined in appendix F.

2.6 Accuracy of Centroiding in Real Data

Several analyses were carried out to determine the errors expected by centroiding. A 4" diameter BK-7 flat polished to $\lambda/2$ P-V was procured and used for fiducial centroiding experiments. An initial experiment was conducted to assess the sizes of fiducials required on flats to be measured with the Wyko 6000 system. The experiments were carried out to establish a baseline and are in no way exhaustive.

2.6.1 Fiducial Sizes

Several circular fiducials were hand-drawn onto the BK-7 flat, ranging between 0.8 and 10 mm in diameter. Ten measurements of the flat were made using the Wyko

6000 and a $\lambda/20$ reference surface without moving the flat or interferometer between measurements. Each of the fiducials 'visible' to the interferometer was centroided using the annulus method. Figure 2.17 shows an image of the part and the resulting phasemap. Statistics were also computed on the number of pixels in each fiducial as shown in table 2.3 At the magnification (zoom) used on the interferometer, each

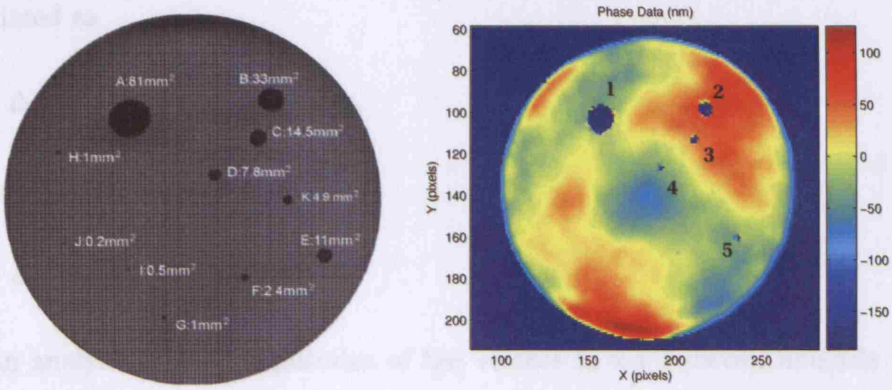


Figure 2.17: Left: image of the BK-7 flat showing the fiducials and their areas. Right: Phasemap of the flat showing the 5 largest fiducials.

Fid No.	1	2	3	4	5
Num Pixels ($\bar{n}$)	162.2	42.7	15.2	5.0	4.2
$\sigma_{\bar{n}}$	0.4	0.3	0.1	0.0	0.3

Table 2.3: Statistics for each of the fiducials shown in figure 2.17.

pixel images an area of 0.48 mm^2 on the flat. The smallest fiducial found, number 5 in figure 2.17, consisted of an average of 4.2 pixels corresponding to a total area of 2.0 mm^2. The actual fiducial was 11 mm^2 in area. Subsequent examination of the original image used to produce figure 2.17 revealed that some of the ink making up the fiducial was very thin around the edges and could be causing the discrepancy. Examination of fiducial number 1 revealed an area of 77 mm^2 as measured by the interferometer and the direct measure was found to be approx 80 mm^2. The ink over

most of fiducial number 1 was found to be consistently thick. Therefore to obtain reliable correspondence between the actual fiducial size and the size as recorded by the interferometer it is a requirement that the fiducial needs to be as uniformly non-reflecting as possible.

The uncertainty in the xy centroids of the fiducials in this experiment, Δ_{xy}, was calculated as

$$\Delta_{xy} = 0.06 \pm 0.02 \text{ pixels},$$

and in z, Δ_z,

$$\Delta_z = 2.04 \pm 0.31 \text{ nm},$$

An analysis of the magnitudes of the vectors in x-y between fiducials was also made. The uncertainty was found to be 0.06 ± 0.02 pixels corresponding to 41 μm on the flat. In this regime, we expect stitching to be accurate to better than 40 μm in x and y. The magnitudes in z had a corresponding uncertainty of 2.90 ± 0.42 nm.

2.6.2 Moving the Part Between Measurements

An experiment was conducted on the same flat with the flat removed from the beam and replaced between each measurement to simulate the different part-camera vector between sub-aperture measurements. The procedure was as follows

1. Position flat in beam and null fringes
2. Measure
3. Remove flat from beam, rotate and replace in a different position
4. Repeat steps 2 & 3 above to give 15 measurements[7].

Three 3 mm diameter circular fiducials were drawn onto the flat and the 11 measurements were analysed; the statistics computed were the centroid of each fiducial

[7]Fractional uncertainty in estimate for σ with 15 measurements is 0.13.

in x, y and z, the number of pixels in each fiducial, the part P-V and the part RMS measurements. Table 2.4 shows the measurement results for the sizes of the fiducials. The fiducial numbers correspond to the fiducials shown in the phasemap in figure 2.16 on page 120. In this experiment it appears that over 11 measurements

Fiducial number	1	2	3	Phasemap	PV (nm)	RMS (nm)
Mean No. pixels	24.82	33.27	31.09	Mean	202.32	34.36
$\sigma_{\text{num pixels}}$	6.59	5.82	5.28	σ	19.8	2.85

Table 2.4: Statistics for each of the 3 fiducials in the second centroid experiment.

the error in the number of pixels assigned as fiducial pixels is approximately ± 6.

Fiducial coordinates were then matched (i.e. stitched) using the least-squares algorithm described in section 4.6 to determine the likely errors to be encountered in stitching. Of the 11 measurement results, the first result was used as the basis for the stitching, that is a transformation was found that matched each data set onto the first. In these analyses the z component of the data is treated separately from the x- and y-components. The phase values represent the deviation of the test surface from the reference surface and are measured in nm. If all the data are scaled into common units then the phase values become totally negligible compared to the real x- and y-coordinates of the data point in question. for example, the x- and y-units might be in mm, but the surface deviation is in nm. If, inside the computer, we compute a vector between a point (x_1, y_1, z_1), and (x_2, y_2, z_2), where the units are all in metres, we may have x and y equal to some mm, and z equal to a few nm. Clearly, we are very insensitive to z as it is approximately 6 orders of magnitude smaller than x and y. We must therefore scale z to have an equal weight on the computation - the easiest way is to make it of the same order of magnitude.

Furthermore, when storing the values in computer memory, the precision of the measurement may be totally lost purely by storing the value as a particular data type such as a single precision *float* - we cannot hope to reliably store and operate on 12.000000123 mm reliably with measurements precise to only a fraction of a mm.

The matching results for the 11 data sets are shown for x and y in figure 2.18 and in figure 2.19 for z. The mean x-y uncertainty in the fiducial positions and its

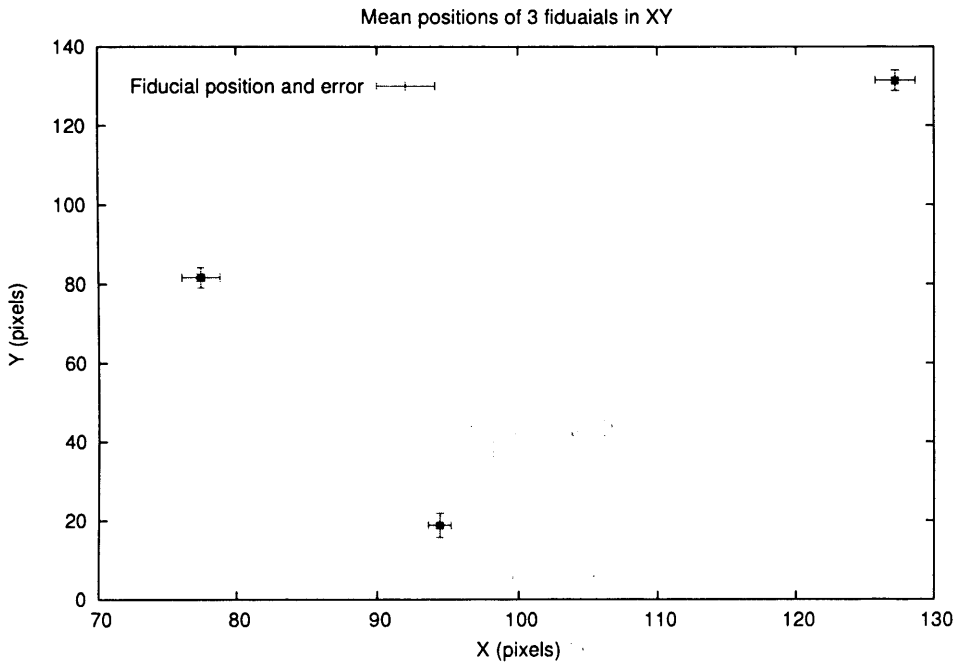


Figure 2.18: Matching 11 sets of 3 points on the same surface in x and y.

corresponding error was found to be:

$$\Delta_{xy} = 1.98 \pm 0.36 \text{ pixels},$$

and that in z, Δ_z,

$$\Delta_z = 2.62 \pm 0.36 \text{ nm}.$$

This z error is a measure of the uncertainty of the length along z of a line joining two parts of a surface together, measured in different frames, or, a measure of the uncertainty of locating a fiducial's z-coordinate in a sub-aperture patch. The z-errors

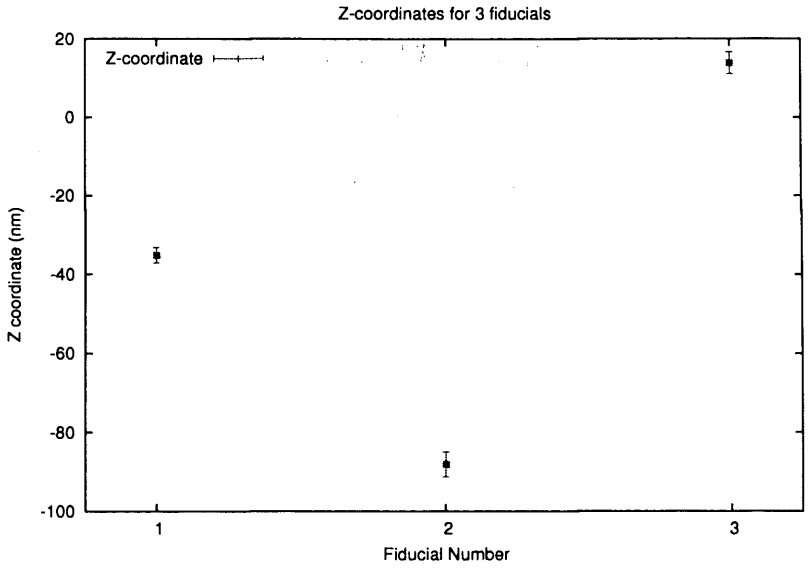


Figure 2.19: Matching 11 sets of 3 points on the same surface in Z.

are consistent with those from the stationary part and the xy-errors are larger. This is expected since each fiducial is imaged with a different camera-target vector for each measurement. The width of 2 pixels here represents 1.4 mm on the surface of the flat - this is a considerable distance and will markedly affect the stitching. A considerable improvement may be made by using a more modern interferometer with a higher resolution and by the maximizing the sub-apertures in the CCD frame if a zoom feature is available.

The general surface statistics show good repeatability in the PV and RMS measurements as expected from a good laboratory interferometer. They are also in good agreement with the statistics from the Wyko acquisition and analysis software.

2.6.3 An Alternative Procedure

The very fact that each pixel in a fiducial in the procedures used above is equally weighted means that only a limited precision is possible when evaluating fiducial centroids. A higher precision might be obtained by consideration of the actual interferogram. For example, centroiding is commonly used in astrometry to give accurate and precise positions of stars on photographic plates or CCD images of

the night sky. These procedures use centroids that are weighted by the intensity of the pixels to achieve high-precision. Recent work by Yano et al [86], has employed a centre-of-mass algorithm to determine a photon weighted mean-position for each star. This position is then revised by comparing the distances between two stars and using a novel least-squares model to derive corrections, where accuracies have been reported at 1/300 pixel.

Centroiding algorithms relying on intensities and position, as referenced above, may be used if a fiducial can be produced with a Gaussian intensity cross-section. However, centroiding with weights is a problem with the raw interferograms because of the fringe patterns that overlay the basic intensity distribution of the returned wavefront. However, most modern interferometers use phase-stepping to give 4 (or 5) interferograms with a phase-shift of $\pi/2$ between each. If we take the mean of these 4 interferograms on a pixel-by-pixel basis, we shall have a distribution corresponding to the mean intensity of the fringes. This will generally approximate a flat field (DC) in the regions of the aperture, but with dark spots corresponding to the fiducials since very little is reflected back into the interferometer. The fiducials can then be centroided on this averaged interferogram using a fully weighted centroid algorithm such as those already in place for astrometry [loc cit.].

Taking the mean of 4 interferograms requires that each one is available from the interferometer. Unfortunately the interferograms are not available from the Fisba and Wyko interferometers and this method could not be used in this work. The author has ascertained that newer interferometers, such as the more recent Fisba model and a model from 4-D Technology Corp. save each interferogram in the raw data file and that this technique may be used in future work to improve the precision of the centroiding.

2.7 Automatic Fiducial Centroiding

The automation of fiducial centroiding is required for an automated machine-based system where it removes the possibility of user errors and allows for increased measurement throughput by minimizing user intervention. To fully automate the task there are many techniques that need to be applied such as image processing and counting methods, for example. Rather than write software from scratch, it is beneficial to utilize existing code wherever possible.

The LabView programming environment was chosen for this automation task for two reasons. Firstly, the software for the swing-arm profilometer is written in LabView and it is natural to write the fiducial centroiding software in the same language for incorporation into the swing-arm code. Secondly, LabView software can be compiled into a fully blown application or Windows library that can be called just like an ordinary Matlab function. Furthermore, LabView contains a very comprehensive set of image processing utilities that are ideally suited to fiducial centroiding. Combining the data acquisition and analysis tools of LabView with the maths engine of Matlab can produce a powerful hybrid software solution to the image acquisition and analysis required for this stitching work.

2.7.1 LabView Centroiding software

There are two approaches to solving the problem of automatic centroiding. The first involves the direct importation of the interferometer binary files into LabView. The data then need to be searched inside the aperture for the `NOT APERTURE` constant whereby the image analysis can take place. This is not the most preferable method since the phasemaps have to be imported into Matlab and Labview. A more convenient method is to pass the phasemap data to Labview from Matlab as the interferometer's binary format has already been converted into a user format in the computer's memory. Furthermore, since the annulus method has been implemented

in Matlab, all that is required from LabView is the automatic centroiding of the fiducials in x and y; Matlab can be used to compute z after the centroids have been passed back by Labview.

Instead of a phasemap, a binary mask is passed to Labview that is composed from the original phasemap. Phase data points are represented by 0, and `NaN` points (fiducials and `NOT APERTURE`) are represented by a 1. A phasemap and corresponding binary mask are shown in figure 2.20. Phase values are not important for the calculation of the x- and y-coordinates of the centroids - it is only important to know which data points form fiducials and it is much easier to operate on a binary array than operate on an array of reals. It is also important to state that at this point the fiducial coordinates are given in terms of the individual phasemaps so the x- and y-coordinates will be given in terms of the axes shown in figure 2.20. Correct scaling of all data sets is carried out at a later stage after the fiducial coordinates have been appended to the phasemap arrays.

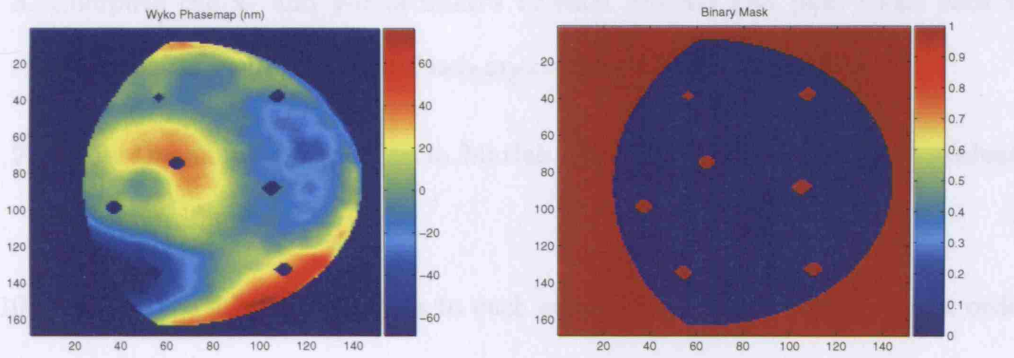


Figure 2.20: Binary mask (right) produced from the phasemap (left).

Procedure

The program must correctly identify full fiducials within the aperture and reject those outside it or those which border the two regions. The task can be broken down into discrete steps:

1. Image the surface with the metrology instrument and load the phase and instrument state data into Matlab.

2. Produce a binary mask of the image where valid aperture data is assigned a value of 0 and all `NaN` pixels are assigned a value of +1.

3. Call the LabView software via Matlab and pass the mask and instrument state data to the LabView centroiding software.

4. Reject fiducials touching the edge of the aperture.

5. Reject objects that are too small to be considered as fiducials.

6. Reject objects whose shape does not conform to that required.

7. Give each remaining fiducial a unique ID containing the global fiducial number and the sub-aperture in which it is to be found.

8. Compute the x- and y-coordinates of each fiducial and pass these back to Matlab along with the global sub-aperture and fiducial numbers.

9. Apply the annulus algorithm in Matlab about the (x, y) centroids to evaluate the z-coordinates.

10. Append the fiducial centroids to each sub-aperture array in the correct order and compute the fiducial matrix.

The LabView *Vision* library contains all the routines required to accomplish the steps 3–8 outlined above. A key feature of the library is that it was designed to create fully autonomous systems for inspection, fault correction and a wide variety of other industrial and scientific image processing tasks, and so is ideally suited to the task here. The steps 3–8 form a set of distinct procedures that have been programmed into LabView to create a machine code library that can be called by Matlab to automatically centroid fiducials. The program code can be found in appendix J.

The production of the binary mask has another advantage. The particle analysis functions in Labview are designed to operate on binary images. Usually, one would filter an image (grayscale or colour) and apply a threshold. Items within the threshold range would be assigned a value of 1 and items outside would be assigned a value of 0. The particle analysis routines can be directly applied to the mask resulting in a very compact and simple Labview program thus avoiding the thresholding requirements.

Figure 2.21 shows the labview GUI for the automatic centroiding of fiducials using the mask shown in figure 2.20, albeit with two additions - an edge fiducial and a small group of pixels near to the centre of field. Contiguous objects whose

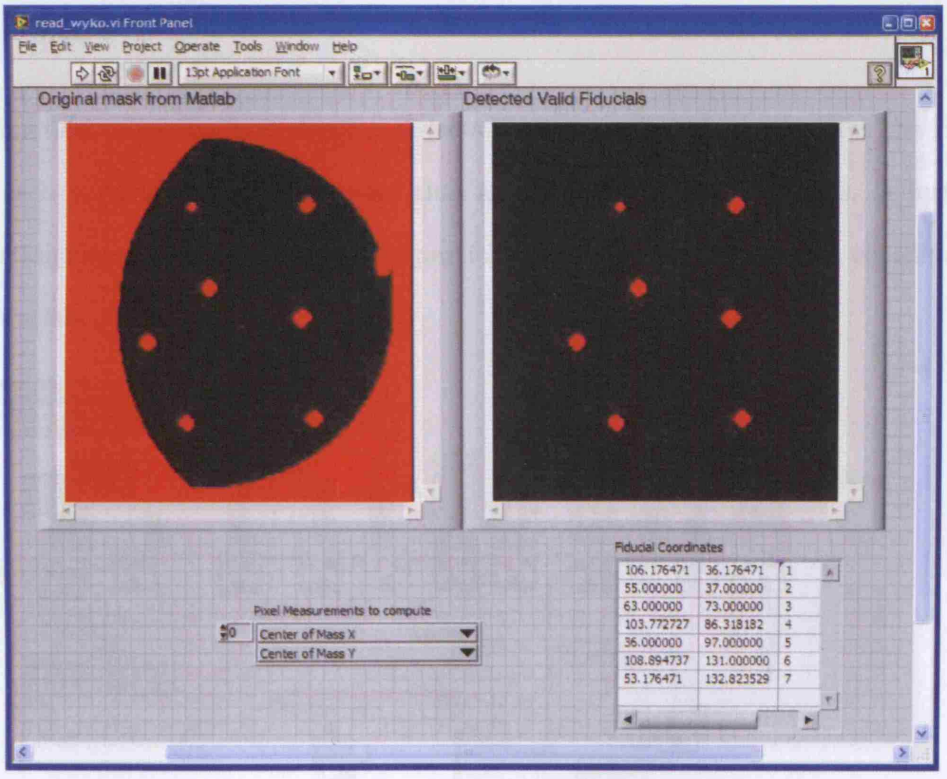


Figure 2.21: LabView centroiding front-end. All invalid data has been rejected leaving 7 valid fiducials and their coordinates. The fiducial touching the edge of the aperture has been rejected and the dead pixels near centre have been rejected - the particle is too small to be a valid fiducial. The whole operation is automatic.

extents touch the edges of the image area have their pixel values set to zero. Next, particles that are too small to be considered are removed. Each object is subjected to an erosion where 3 layers of pixels are set to zero from the perimeter of the object. If nothing is left after the erosion then the particle is deemed too small and all the values are left at zero. If some non-zero pixels remain then all the original pixel values are retained for that object. Now the image field contains a series of objects none of which touch the edges and all are above a certain size. The *Heywood Circularity Factor* (CF) now provides a useful measure of the remaining objects. This is defined as the perimeter of an object divided by the circumference of a circle having the same area as the object:

$$\mathrm{HCF} = \frac{P}{2\sqrt{\pi A}}. \tag{2.13}$$

A function to remove non-circular objects is then applied that calculates the HCF for each particle and removes those that are not within required limits. A further function then centroids each unique particle (fiducial) and outputs the coordinates to Matlab.

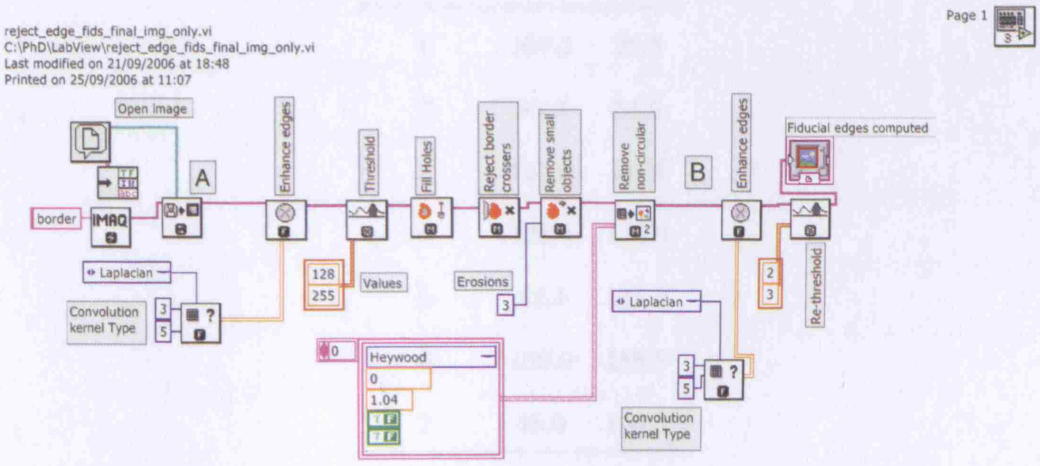


Figure 2.22: The Labview visual code to automate the fiducial centroiding.

To illustrate the software functionality, a test image was generated using a simple imaging system composed of a commercial CCD camera and lens combination. An

image was made of a series of spots hand-drawn onto a glass plate. This image was then loaded by Matlab as an array, passed to the LabView function where automatic centroiding took place before the centroids were finally passed back to Matlab. Figures 2.23 and 2.24 show the LabView processing stages and the Matlab GUI respectively. The centroids for figure 2.24 are indicated in table 2.5

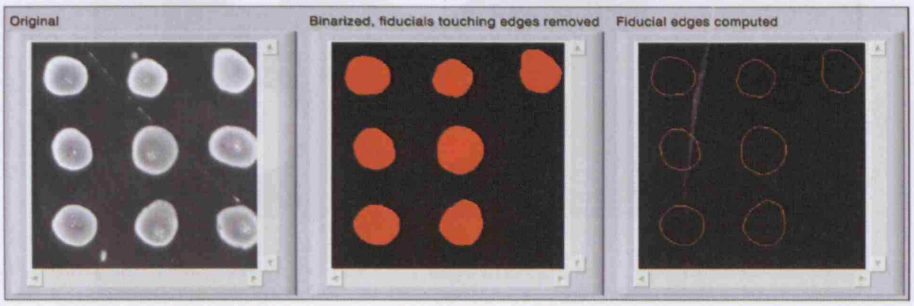


Figure 2.23: Labview processing stages for the software test. In this case, the raw image data was passed to LabView before thresholding and automatic centroiding. A final extra stage is also shown here where a mask containing the layer of pixels immediately surrounding the fiducial is computed.

Fiducial	x	y
1	209.2	29.5
2	35.2	34.5
3	121.0	36.8
4	128.4	112.0
5	42.4	111.2
6	130.6	188.8
7	45.0	190.3

Table 2.5: Centroids for the fiducials indicated in figure 2.24. Centroids are listed using a raster scan starting at the top right and proceeding right-to-left and down.

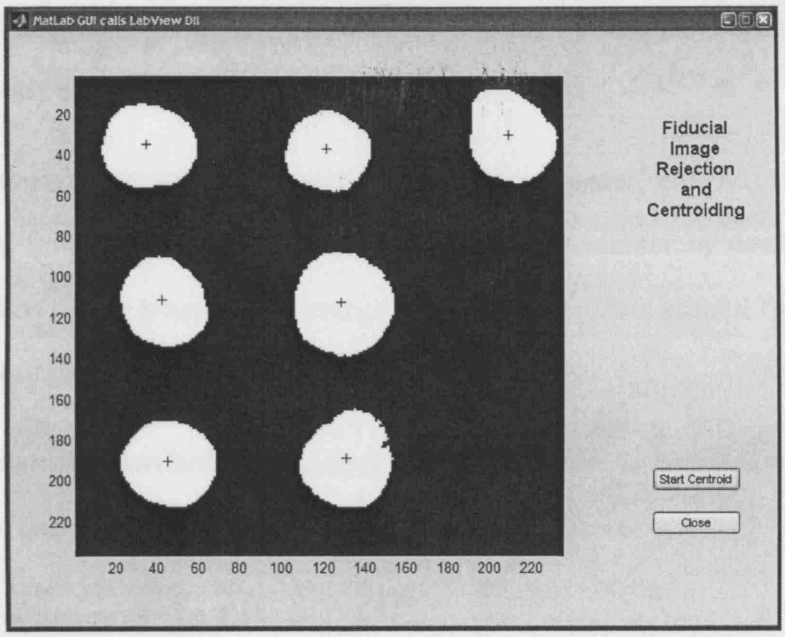


Figure 2.24: Matlab centroiding screen after LabView image processing. The blue crosses represent the fiducial centroids as computed automatically by LabView. Table 2.5 gives the coordinates of the fiducials shown here.

2.8 Summary and Concluding Remarks

In this chapter the following important points have been made:

1. The stitching algorithm described in this work will stitch optical metrology data sets by matching fiducial points common to the overlapping areas of adjacent sub-aperture measurement areas. The stitching transformations will be derived by a mathematical optimization to the least-squares problem expressed by equation 2.8. The equation will be extended to include minimizing the residual for all sub-apertures simultaneously.

2. The surface metrology will be performed by a phase-shifting optical interferometer. This constitutes an important work-around due to the non-delivery of the OMAM phase diversity wavefront sensor.

3. The surfaces whose form is to be measured will be fiducialized by the application of ink spots from felt-tipped marker pens. This physical surface

fiducialization was set down as **a requirement by the project partners** at the start of the project.

4. A successful method has been demonstrated to locate (centroid) the positions of the fiducials in the output data of the interferometer by consideration of the data points both constituting, and also those points around the fiducial in the recovered phasemap.

5. An algorithm has been demonstrated for the automatic location and centroiding of the fiducials in the phasemap data.

6. The accuracy of the fiducial centroiding has been shown to be 1.98 ± 0.36 pixels in x and y, and 2.62 ± 0.36 nm in z in the circumstances described in section 2.6.2.

The next chapter follows the development of the stitching algorithm in terms of the types of transformations that are required and the prior knowledge gained from fiducialization. This is followed by a chapter detailing the numerical optimization and ancillary procedures used to find the 'best' stitching parameters.

CHAPTER 3

Geometric Transformations for Stitching

"Prudens quaestio dimidium scientiae."

(Half of science is asking the right questions.)

– Roger Bacon, Franciscan friar and philosopher (c.1214–1294)

3.1 Applying Transformations

This chapter is concerned with developing the mathematical abstractions and concepts used in the formulation of the stitching algorithm. The work then proceeds to set-out the techniques developed to access the fiducial centroid data gleaned from chapter 2.

From the beginning it was desirable to develop an approach to the software solution that would be extensible such that no re-writing of code would be required to switch from stitching a given number of sub-apertures to a larger number. The main stitching algorithm developed in this chapter and the next is therefore applicable

to stitching any number of sub-apertures in any overlapping configuration where the only limitation is the amount of memory available in the computer running the algorithm. The software automatically determines the pairings of the fiducials and thus the configuration of the sub-apertures, regardless of their quantity. Previous literature on sub-aperture stitching has tended to concentrate on stitching simple aperture outlines such as spheres or rectangles and often using rectangular arrays of sub-apertures to cover them [66, 64, 87, 71, 88, 89]. To be completely general where a single surface is concerned a stitching system should be able to stitch sub-aperture measurements on any of the surfaces shown in figure 3.1. The surfaces

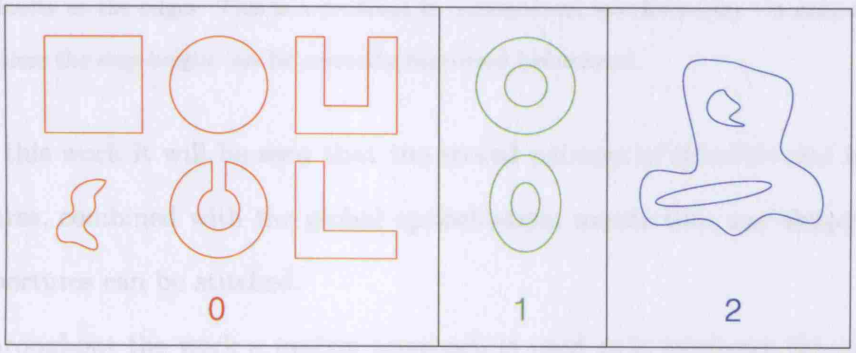


Figure 3.1: Examples of surface shapes that could stitched by sub-aperture stitching.

labelled '0' in red are have a genus of 0 and are topologically equivalent as they can be continuously deformed into each other. Roughly speaking the genus is equal to the number of holes in the surface [90]. The same applies for the surfaces in green and blue with genuses of 1 and 2 respectively. A Cassegrain telescope mirror has a genus of 1 and it is therefore desirable that the stitching software can stitch surfaces of any genus. The only constraint to consider is that the parts of the surface to be measured are smooth, i.e. continuous in first derivative. For example, the surface existing inside the red boundary of the square in figure 3.1, may take the form of a plane or a sphere for example, and this applies to all the shapes shown. If the shape is extruded out of the paper, one might consider measuring the ends of the 'U' as

shown in figure 3.2. This is not possible, for although the ends are connected to the rest of the part, there is a discontinuity between the tops of the 'u'.

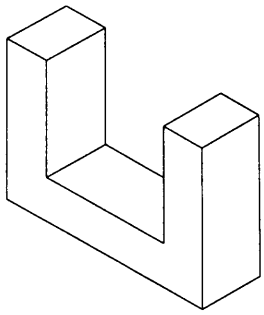


Figure 3.2: The ends of the 'U' cannot be measured simultaneously or by stitching due to the discontinuity at the edges. This is a problem in conventional interferometry - it cannot work over steps unless the step-height can be correctly measured beforehand.

In this work it will be seen that the stored pairings of fiducials and hence sub-apertures, combined with the global optimization, means that any shaped array of sub-apertures can be stitched.

Throughout the work a matrix approach is used as it expresses the operational aspects in the most convenient way.

3.1.1 The Requirements for Transformations.

In chapter 2, a least-squares condition for stitching two sub-apertures was set forth in equation 2.5. The equation is valid when real tip/tilt exists in the data as distinct from rotation, and was used by Otsubo et al. [loc cit.] to connect interferograms before phase recovery. Tip and tilt expressed in this way are shearing operations and they scale the data in z and therefore are not shape preserving, but they are valid for real wavefront tip and tilt[1]. If the tip/tilt angle is small then an approximation can be made using a rotation since $\tan\theta \simeq \theta$ for small θ. This approximation has

[1]Imagine a collimated beam with a circular footprint incident upon a plane surface tilted at an angle to a plane parallel to the beam wavefronts. The beam footprint will take the form of an ellipse, not a circle, i.e. it has been sheared.

been used in past work by Cheng et al. [87] for stitching, on the basis that if the interferograms are nulled before measurement, the angle is by definition very small indeed. In their work they simulated two sub-apertures from one large one. To one sub-aperture they added a piston and tip/tilt terms in the form of rigid rotations about the lines x2 and y2 as shown in figure 3.3:

> "The coordinate of sub-aperture 2 is rotated some relative to that of sub-aperture 1 around the x2 and y2 axes to simulate the relative tilts B and C between sub-apertures." [87]

Cheng's perturbing (rotating) formula is therefore valid only for small tip/tilt angles since their stitching function corrects tip and tilt in the form Ax and By which is equivalent to shearing parallel to z.

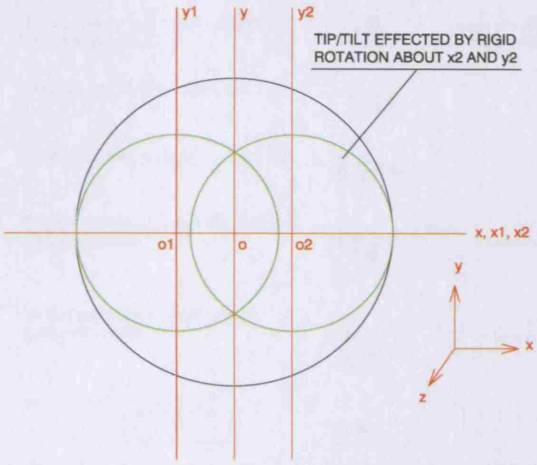


Figure 3.3: Sub-aperture stitching scheme redrawn from [1]. The authors simulated some data and applied rotations to the data around x2 and y2 to simulate tip/tilt. They then re-stitched the data using a shearing operation on the basis that the tip/tilt angle is very small.

Output Data

Most modern interferometers are capable of producing an output that directly gives the form, or the error from the reference surface, of the surface under test. It is a fact that a raw phasemap will have incorrect scaling, in the radial and z-directions and also in the lateral directions. For example, for a surface measured in reflection, the

raw phase values have to be multiplied by 0.5 since, relative to the reference surface, the light has travelled into a 'hole' and back out, twice the depth of the hole. Most interferometry analysis packages come with controls for the type of test conducted, e.g., single or double reflection etc. The well known Fisba μShape software contains a configuration screen to allow the user to enter the type of optical test to facilitate the correct scaling of the phase data. An example screen is shown in figure 3.4.

More generally, a scaling coefficient sometimes called a wedge factor can be specified in the interferometer's software to correct matters in relation to the test configuration.

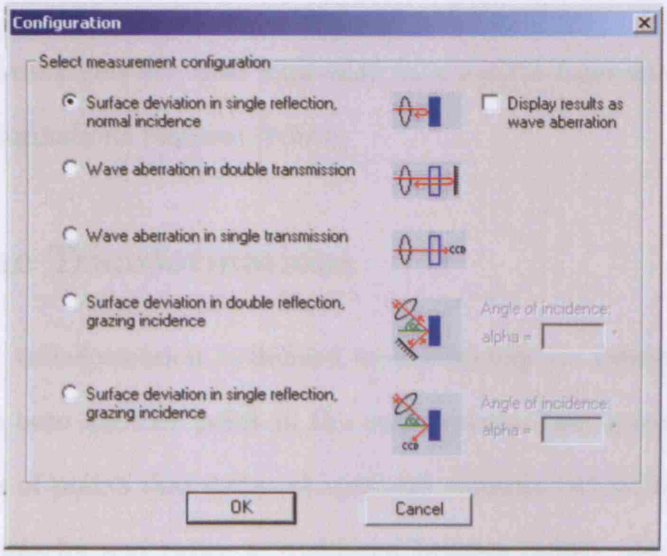


Figure 3.4: Fisba μShape test set-up configuration screen.

Post-output, phase or surface data for each sub-aperture will be tip/tilt corrected and will give the surface form or the error from the datum surface.

Several overlapping sub-aperture phasemaps in this form can be stitched together using **rigid rotations and translations**, along with the tip/tilt correction. Each sub-aperture measurement will have its own tip-tilt correction, but this will generally not be equal to that for the entire stitched data set. Furthermore, we imagine a general measurement platform where any sub-aperture may be measured in any

orientation and the general direction corresponding to 'up' in the interferometer data of a given sub-aperture will not correspond to 'up' in the part orientation. We therefore require at least:

- A polar rotation of the phasemap about the z-axis.
- Tip/Tilt correction terms for the phasemap.
- A general (x, y, z) translation vector.
- We may also include a wedge factor to account for the optical set-up if this cannot be applied in the interferometer software.

The transformations are best expressed in a matrix form and a review of the types of transformations required follows.

3.2 Affine Transformations

A geometrical transformation is defined as a mapping (or function) that maps a point in space onto another point in the same space. The transformation can be applied to sets of points that define shapes and volumes. An affine transformation preserves co-linearity and ratios of distances between points. Contraction, expansion, dilation, reflection, shear and translation are all affine, as are combinations of these. However, if we wish to stitch two sub-aperture patches whose form is known, we cannot arbitrarily scale (expand/contract) or reflect. This would otherwise alter the form of our patches. The transformations must be restricted to rigid translations, rotations and tip/tilt corrections in the form of shear, with the occasional post-stitching scaling.

This work considers affine transformations, or mappings, of the forms:

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^n, \qquad n = 3 \quad \text{or} \quad 4 \tag{3.1}$$

where,

$$F(\mathbf{p}) = \mathbf{A}\mathbf{p} + \mathbf{q}. \tag{3.2}$$

Here, $\mathbf{A}$ is a linear matrix transformation in $\mathbb{R}^n$ and $\mathbf{p}$ and $\mathbf{q} \in \mathbb{R}^n$. An interesting point to note is that if $\det(\mathbf{A}) > 0$, the transformation is orientation preserving, otherwise if $\det(\mathbf{A}) < 0$ the transformation is orientation reversing, see for example, Weisstein [91]. The transformations for which $\det(\mathbf{A}) < 0$ are called rotoinversions and result in the inversion of the data. For stitching metrology data, only transformations for which $\det(\mathbf{A}) = 1$ can be considered unless some small scaling allowance is made for aspheric parts where $\det(\mathbf{A}) \neq 1$ may be permitted. A scaling may be used on aspheric parts to allow for changes in the radius of curvature over the part's surface. The change in radius of curvature of the part alters the scale of the measurement area on the interferometer CCD as different sized areas are measured. Therefore, in future work, the algorithm may be modified to allow an extra scaling parameter to be included in the optimization routine.

The conventions used for describing matrix transformations vary, but in this work all transformations are based on post-multiplying by a column vector, as indicated in equation 3.2, since it will be seen that the data are stored as 3×n, or 4×n arrays in computer memory.

We wish to combine arbitrary polar rotation (about the z-axis), translation and the tip/tilt correction to produce a generalized transformation matrix that can be used by the stitching software. Consider the matrix equations 3.3, 3.4 and 3.5:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \quad \mathbf{p}' = \mathbf{p} + \mathbf{t}, \tag{3.3}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{p}' = \mathbf{R}\mathbf{p}, \tag{3.4}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{p}' = \mathbf{S}\mathbf{p}. \tag{3.5}$$

In $\mathbb{R}^2$ they provide translation, rotation by angle θ about the z-axis and scaling about the origin respectively. The translations are also affine and are three types used in this work. An interesting point to note is that $\mathbf{R}$, in equation 3.4, is an orthogonal transformation with $\det(\mathbf{R}) = 1$. Also, in equation 3.5, $\det(\mathbf{S}) = s_x s_y$ is the area scale factor and s_x and s_y are the scale factors in x and y respectively.

A single $\mathbb{R}^3$ compound transformation matrix that accomplishes at least translation and rotation in a single step is required. In equations 3.3, 3.4 and 3.5, the translation is accomplished by addition and the others are multiplications. We can express them all by multiplication by the methods of homogeneous coordinates.

3.2.1 Homogeneous Coordinates

Homogeneous coordinates, defined below, are used very frequently in computer graphics and in particular, in Computer Aided Design (CAD) programs. They are also used in machine tool technology since in this respect, one might have a CMM or other measuring tool that has linear slides and rotation stages. The coordinates of a measurement tip might need to be expressed in Cartesian coordinates from the machine's own coordinate system composed of the machines motion axes. By using homogeneous coordinates one can express the whole coordinate transformation in terms of a single matrix, thus avoiding the time spent in multiplying the machine coordinates by several matrices.

Homogeneous coordinates are also used in chapter 5 where a coordinate transformation is derived to transform the swing-arm profilometer coordinates to Cartesian space in the form of a single transformation matrix.

Foley provides an excellent treatment of homogeneous coordinates from the point of view of computer analysis. *We add a third coordinate to each point to use homogeneous coordinates in* $\mathbb{R}^2$ [92]. Thus, (x, y) becomes (x, y, W). Each point now has infinitely many homogeneous coordinate representations and any two coordinates

that are multiples of each other express the same point. A point is *homogenized* when $W \neq 0$ as $(x/W, y/W, 1)$. Now, $(x/W, y/W)$ are the Cartesian coordinates of the point. We usually choose $W = 1$ and therefore homogenized points form the plane defined by $W = 1$ in (x, y, W) space.

We can now expand equations 3.3, 3.4 and 3.5 to take advantage of homogeneous coordinates and begin to form *compound* matrices.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ 1 \end{bmatrix}, \quad \text{or} \quad \mathbf{p}' = \mathbf{T}\mathbf{p}, \tag{3.6}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad \text{or} \quad \mathbf{p}' = \mathbf{R}\mathbf{p}, \tag{3.7}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad \text{or} \quad \mathbf{p}' = \mathbf{S}\mathbf{p}. \tag{3.8}$$

These can be shown to be equal to equations 3.3, 3.4 and 3.5 after multiplication and then by considering the x and y coordinates, i.e. homogenizing. To construct a compound rotation and translation matrix we simply multiply the matrices together in the appropriate order. For example, we may wish to rotate all points by an angle θ in the x-y plane about a point (p_x, p_y). To achieve this we first translate by $(-p_x, -p_y)$, then rotate about the origin and then re-translate back by (p_x, p_y), as illustrated in figure 3.5.

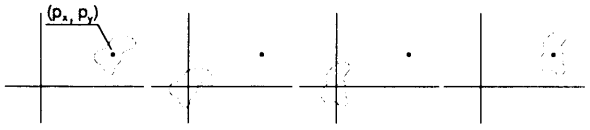


Figure 3.5: Composting transformations: rotation about an arbitrary point in $\mathbb{R}^2$ is accomplished first by a translation, then by a rotation and then by a final translation.

The $\mathbb{R}^2$ compound transformation matrix is simply composed by multiplying the individual matrices together:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.9}$$

$$\mathbf{T} = \begin{bmatrix} \cos\theta & -\sin\theta & t_x(1-\cos\theta) + t_y \sin\theta \\ \sin\theta & \cos\theta & t_y(1-\cos\theta) + t_x \sin\theta \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.10}$$

Consider the upper left 2×2 submatrix in equation 3.10. Each of the 2×1 row vectors forming the sub-matrix is a unit vector, since $\sin^2\theta + \cos^2\theta = 1$, and they are orthogonal to each other since their dot product evaluates to zero. This matrix is termed *special orthogonal* and has a determinant equal to 1. In general a transformation matrix of the form:

$$\mathbf{T} = \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.11}$$

where the upper left 2×2 submatrix is orthogonal preserves angles and lengths, and is therefore affine. These transformations are linear conformal, or rigid-body transformations and form a subset of the more general affine group.

The use of homogeneous coordinates can be extended to $\mathbb{R}^3$ by the addition of a 4th coordinate to the $\mathbb{R}^3$ transformation matrices. A general point $\mathbf{p}$ becomes $(x, y, z, 1)^{\mathrm{T}}$ and a general transformation matrix $\mathbf{T}$ becomes

$$\mathbf{T} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.12}$$

3.2.2 Application to Data Stitching

Figure 2.3 on page 93 shows a sequence for stitching two sub-aperture measurements using affine transformations to eliminate the distances between fiducial pairs one-by-one. Figure 3.6 on page 146 shows a similar figure for exact stitching by illustrating a possible sequence of analytical procedures. Here it is assumed that the surface patches stitch exactly and that direct vector operations can be applied after computation of angles and distances. The figure illustrates stitching by way of rotations and translations for a theoretical data set. Correcting real tip/tilt is dealt with later.

The act of applying several different transformations in sequence is equivalent to composting a single transformation matrix to achieve all the steps at once. The order of operations is not important, but there are several steps that are necessary. Here is one such set of steps on the perfect data set:

Translation: A $\mathbb{R}^3$ translation is required to move the second patch to match at some point with the first. At least one point in the overlap area must match, but we can envisage scenarios where a line of points match or where all overlap points match - this would occur if there was no tip/tilt or rotation between our idealized sub-apertures. In figure 3.6, the translation $A' \rightarrow A$ is made first along the vector $\mathbf{t}$.

Rotation: Up to three $\mathbb{R}^3$ rotations may be required in a Cartesian system. For example, if one fiducial is matched by translation as above then two **general** rotations are required to stitch the data perfectly, general in the sense that they are not parallel with any of the coordinate axes. In the example, one rotation will be about the line through A along the cross product of $\mathbf{b}' \times \mathbf{b}$ and the second will be about the line along the vector $\mathbf{b}$. The rotations about $\mathbf{b}' \times \mathbf{b}$ or about $\mathbf{b}$ can be achieved in several ways but generally involve two rotations in a $\mathbb{R}^3$ Cartesian system along vectors parallel to the axes. These

steps and the named vectors are shown in figure 3.6.

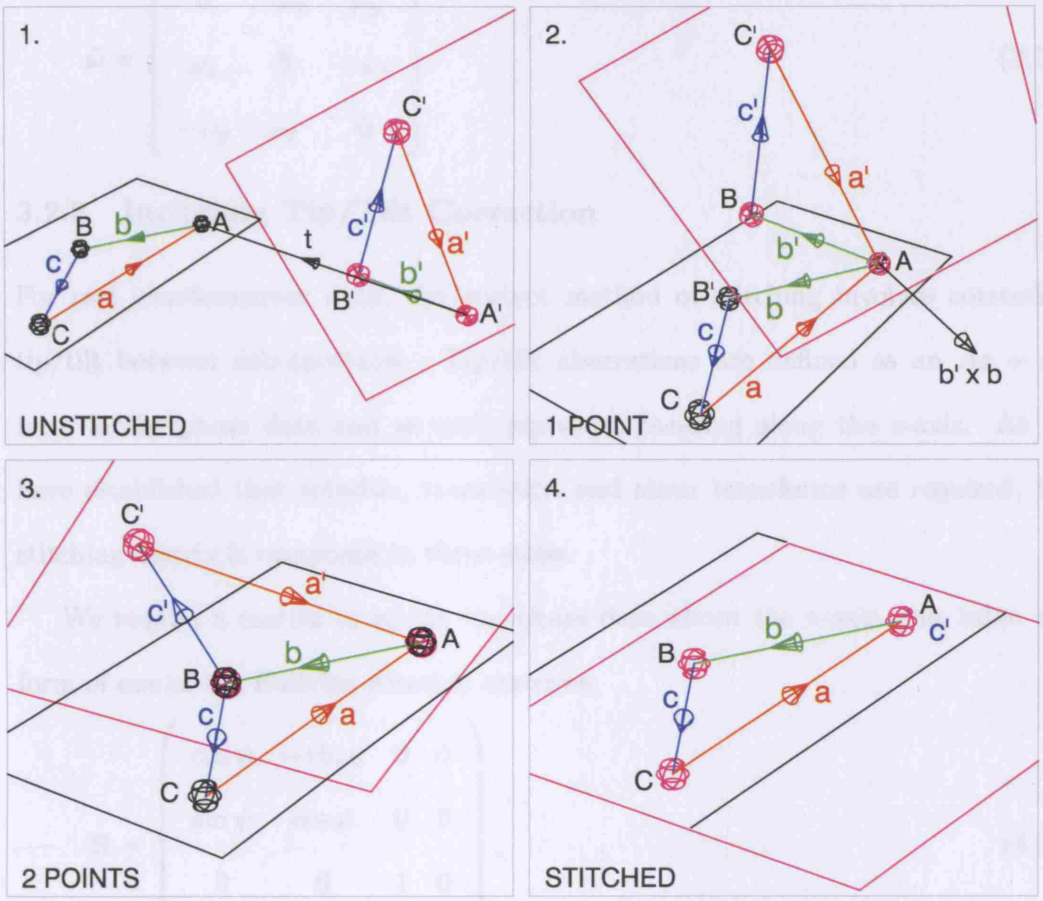


Figure 3.6: Stitching perfect data. 1. A translation $A' \rightarrow A$ is made on the points in the sub-aperture to stitch. 2. A rotation of points is made about $\mathbf{b}' \times \mathbf{b}$ to make B and B′ coincident. 3. A final rotation about $\mathbf{b}$ is made to make C and C′ (and therefore all points) coincident. The whole process above assumes that the data are perfect i.e., the points A, B and C are perfectly congruent.

The rotation matrices for this example may be computed in a variety of ways, but a common method is to use Rodrigues' special orthogonal rotation matrix [93]. The matrix, to rotate by an angle θ about a fixed axis specified by the unit vector $\hat{\omega} = (\omega_x, \omega_y, \omega_z)$ is given by:

$$e^{\tilde{\omega}\theta} = \mathbf{I} + \tilde{\omega}\sin\theta + \tilde{\omega}^2(1 - \cos\theta), \tag{3.13}$$

where $\mathbf{I}$ is the identity matrix and $\tilde{\omega}$ is the antisymmetric matrix given by

$$\tilde{\omega} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}. \tag{3.14}$$

3.2.3 Including Tip/Tilt Correction

For real interferometer data, the correct method of stitching involves correcting tip/tilt between sub-apertures. Tip/tilt aberrations are defined as an $Ax + By$ term in the phase data and as such represent shearing along the z-axis. As we have established that rotation, translation and shear transforms are required, the stitching matrix is composed in three steps.

We require a matrix to rotate the phase data about the z-axis, this takes the form of one of the Eulerian rotation matrices:

$$\mathbf{R} = \begin{pmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.15}$$

which is a homogeneous matrix for $\mathbb{R}^3$ and can be constructed by computing the positions of the $\mathbb{R}^3$ basis vectors after the counterclockwise rotation by angle θ about the z-axis. Alternatively, we compute the matrix using Rodrigues' method in equation 3.13 with the unit vector $(0\ 0\ 1)^T$.

Now, the translation matrix is given by

$$\mathbf{T}_{xyz} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.16}$$

where t_x, t_y and t_z are the components of the translation.

The tip/tilt correction matrix is given by:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ A & B & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.17}$$

where A is the tip coefficient and B is the tilt coefficient. The compound matrix is found my multiplying the matrices $\mathbf{AT_{xyz}R}$, to yield:

$$\mathbf{T} = \begin{pmatrix} \cos\psi & -\sin\psi & 0 & t_x\cos\psi - t_y\sin\psi \\ \sin\psi & \cos\psi & 0 & t_x\sin\psi + t_y\cos\psi \\ A & B & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.18}$$

$\mathbf{T}$ defines the main transformation matrix used in this work. Relative piston between sub-apertures is accounted for in the translation matrix $\mathbf{T_{xyz}}$, but an alternative would be to omit the z-translation here and include it in the matrix $\mathbf{A}$ as the coefficient C at $\mathbf{A}_{3,4}$, to give the new, corrected z-coordinate as:

$$z' = Ax + By + z + C. \tag{3.19}$$

3.2.4 Stitching Parameters

The definition of $\mathbf{T}$ for the matrix transformation indicates that 6 parameters are required for each sub-aperture: tip, tilt, a polar rotation and the three linear translation parameters. Let the transformation parameters for the i^{th} sub-aperture be denoted by $\mathbf{t}_i$,

$$\mathbf{t}_i = \begin{bmatrix} A_i & B_i & \psi_i & tx_i & ty_i & tz_i \end{bmatrix}^T \tag{3.20}$$

3.2.5 Euler Rotations

It was mentioned in section 3.1.1, that an approximation can be made when the tip and tilt terms are small. We may approximate the tip/tilt by a rotation. The

resulting transformation is more complicated - and therefore less efficient and not used in this work - but it may find uses elsewhere in stitching other types of data, and is therefore included as an aside.

Euler's rotation theorem provides a more convenient means than Rodrigues' formula for describing an arbitrary rotation with just three parameters. Indeed, the full position of a body in $\mathbb{R}^3$ can be completely described by the three coordinates of its centre of mass and three angles which determine its orientation relative to the coordinate axes. The notation used here differs slightly from Landau and Lifshitz [94], and Weisstein [95] but the overall result is equivalent.

Let the rotations about the x-, y- and z-axes be **B**, **C** and **D** respectively. Then a general rotation matrix can be written as $\mathbf{A} = \mathbf{BCD}$ and let the angles used in the matrices be ϕ, θ and ψ respectively[2].

For the rigid rotation transformations (not used in this work), the transformation parameters the of ith sub-aperture become,

$$\mathbf{t}_i = \begin{bmatrix} \phi_i & \theta_i & \psi_i & tx_i & ty_i & tz_i \end{bmatrix}^T, \tag{3.21}$$

and the required compound matrix, $\mathbf{T} = \mathbf{BCDT}_{xyz}$ evaluates to:

$$\mathbf{T} = \begin{pmatrix} \cos\psi\cos\phi - \sin\psi\cos\theta\sin\phi & -\cos\psi\sin\phi - \sin\psi\cos\theta\cos\phi & \sin\psi\sin\theta & t_x \\ \sin\psi\cos\phi + \cos\psi\cos\theta\sin\phi & -\sin\phi\sin\psi + \cos\psi\cos\theta\cos\phi & -\cos\psi\sin\theta & t_y \\ \sin\theta\sin\phi & \sin\theta\cos\phi & \cos\theta & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.22}$$

T can be thought of simply as providing a rotation about the origin followed by a translation or as a rotation about an arbitrary point, depending upon the translation used. For a given upper-left 3×3 rotation submatrix the 3×1 vector in the upper-right can be varied to move all points to anywhere in $\mathbb{R}^3$. This stitching matrix could be used with the optimizer software developed in the next chapter to stitch data other than optical metrology data.

[2]Several conventions are used in textbooks for naming angles, and indeed, here, we rotate about the x-axis , then the y-axis and finally about the z-axis. In Landau [loc cit.] the rotations proceed about the z-axis, the x-axis and then finally about the z-axis again. These are equivalent: the resulting compound matrix has the same elements.

3.2.6 Global Stitching

Let the transformation matrix for the i^{th} sub-aperture be denoted by $\mathbf{T}_i$, as given by equation 3.18. Once the optimum $\mathbf{T}_i$s have been found they can be applied to the sub-aperture data to perform the stitching. Finding the $\mathbf{T}_i$s is the subject of the next chapter, but before the method used is described, the other parts of the stitching algorithm and how the sub-apertures are stitched are first considered.

It is important to realise that the stitching can be carried out in several ways. One sub-aperture can be stitched onto another, and then further sub-apertures can be added one-by-one. However, whilst the least-squares solution can be found between any two sub-apertures, in this way, the overall stitched result of n sub-apertures will not necessarily be stitched in a global least-squares sense. Therefore, the work in the following sections and chapter, *seeks to illustrate a practical method of minimizing the sum of the squared distances between matching fiducials for all common areas simultaneously*. This aspect of stitching has only been considered in the literature by Otsubo and Fleig. [1, 65] since it is sometimes easy to assume implicitly that this will be the case for stitching.

Stitching $n \times 1$ sub-apertures is a special case because each sub-aperture overlaps at most, only two others. Furthermore, no more than two sub-apertures share a common overlap area. We can stitch two sub-apertures together first and then stitch another to the second sub-aperture because the third sub-aperture does not overlap the first. In general, sub-aperture number $n + 1$ does not overlap any sub-aperture number less than, n, in a $n \times 1$ scheme. Figure 3.7 illustrates this concept. Sub-apertures therefore must only match their sole or two neighbours in a least-squares sense, rather than matching at least two other sub-apertures, all with a common overlap area. An algorithm that takes advantage of this special case is presented in section 4.6.

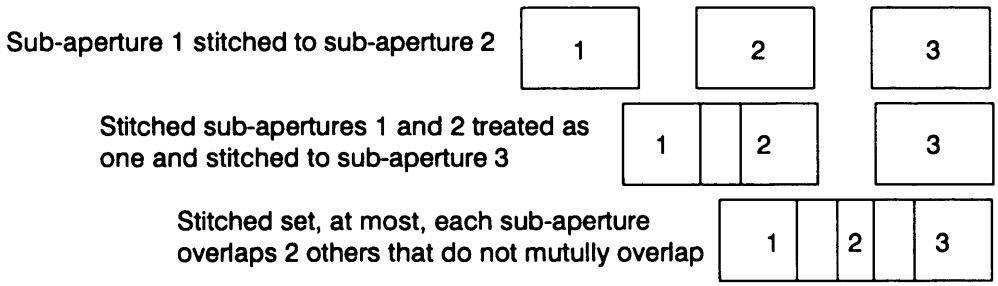


Figure 3.7: Stitching $n \times 1$ sub-apertures. Each sub-aperture only overlaps, at most, two other sub-apertures, and each overlap area is unique to only one sub-aperture pair. Therefore stitching can proceed by stitching sub-aperture 1 to sub-aperture 2, followed by stitching this stitched set as a whole unit to sub-aperture 3 and so on.

3.3 Procedural Stitching Algorithm

Once the stitching parameters, $\mathbf{t}_i$ are known, the matrix in equation 3.18 can be applied to the points in a sub-aperture. It is worth restating here that during the process of finding the stitching parameters, only the coordinates of the fiducials are transformed to save time. Section 2.5.2 dealt with appending the fiducial coordinates onto the end of the sub-aperture arrays, and combining this with the theory of homogeneous coordinates we arrive at the form of the sub-aperture array ready for the stitching, complete with the fourth dimension added:

$$\mathbf{S}_n = \begin{bmatrix} x_1 & x_2 & \dots & \dots & \dots & \dots & x_i & fx_1 & fx_2 & \dots & fx_m \\ y_1 & y_2 & \dots & \dots & \dots & \dots & y_i & fy_1 & fy_2 & \dots & fy_m \\ z_1 & z_2 & \dots & \dots & \dots & \dots & z_i & fz_1 & fz_2 & \dots & fz_m \\ 1 & 1 & \dots & \dots & \dots & \dots & 1 & 1 & 1 & \dots & 1 \end{bmatrix}. \tag{3.23}$$

Here i denotes the i$^{\text{th}}$ point in that sub-aperture, $(x_i,\ y_i,\ z_i,\ 1)^T$ are the coordinates of the i$^{\text{th}}$ homogeneous data point and $(fx_m,\ fy_m,\ fz_m,\ 1)^T$ is the centroid of the m$^{\text{th}}$ fiducial point. The advantage of appending the fiducials to the array becomes clear when one considers the technicalities of accessing array data in Matlab and the use of the fiducial matrix in section 3.3.4.

3.3.1 Expansion of the Flowchart

The basic stitching flowchart from figure 2.2 on page 89 can now be expanded and revised to include the developments described so far. Figure 3.8 shows the revised scheme incorporating blocks for surface fiducialization, least-squares summing and centroiding, amongst others. The centroiding block is dealt with in section 2.5.2 and

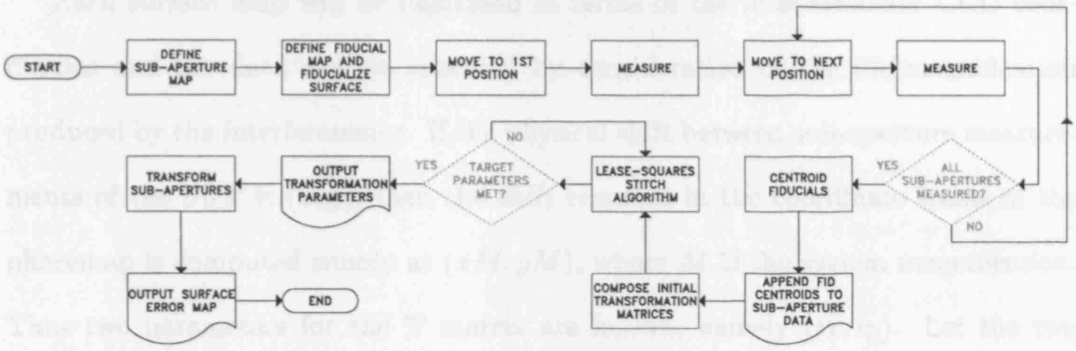


Figure 3.8: Second-generation flowchart for the fledgling stitching algorithm.

this also covers appending the fiducial coordinates to the end of the sub-aperture arrays. Composition of the transformation matrices has been shown in section 3.2.3. The remainder of this chapter deals with estimating the parameters in the $\mathbf{t}_i$ and the final transformation of the sub-apertures once the optimal transformations have been found. A further subsection deals with the method used to determine fiducial pairings using a matrix that has been named the *fiducial matrix*.

3.3.2 Estimating the Initial Transformation Parameters

The first transformation matrices are computed based on the sub-aperture configuration at the time of measurement and this is most easily illustrated by way of an example.

Consider the stitching of two sub-apertures measured on a flat with a nominal displacement of (x, y) between measurements. The data from the interferometer will contain the difference between the reflected wavefront from the surface under test

and the wavefront reflected from the reference surface (the datum). After processing, each phasemap will be located in a rectangular array of size $(i \times j)$, and in general, the interferometer will produce equal array sizes for each phasemap, the aperture and fiducial data occupy various proportions of this whole array whose size is usually dictated by the properties of the imaging device.

Each surface map will be expressed in terms of the interferometer CCD coordinates and the data can be re-scaled by consideration of the (de)magnification produced by the interferometer. If the physical shift between sub-aperture measurements of the SUT is (x, y), then the shift required in the coordinate frame of the phasemap is computed simply as (xM, yM), where M is the system magnification. Thus two parameters for the $\mathbf{T}$ matrix are known, namely (tx, ty). Let the two phasemaps have a mean value of zero and we may assert that the piston term between sub-aperture measurements is small and we therefore set tz to equal zero. The tip/tilt parameters, A and B, are unknown but generally small and in this instance and we therefore set them to zero. If the interferometer or part was rotated between the measurements, we simply insert this rotation angle, ψ, into the $\mathbf{t}_i$.

In a more complex scenario where the part or the interferometer is mounted on a translation stage we can derive the parameters from the readout sensors on each of the machines axes. These values are then the initial guess or seed values for the stitching algorithm optimizer. The details of the numerical optimization are given in section 4.4, but the steps to be taken after the least-squares condition has been achieved follow here as they are all matrix operations that fit with the content of the other parts of this chapter.

3.3.3 Transforming the Sub-Apertures

Let the set of optimum transformation matrices be:

$$\{\mathbf{T}\} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_n\}, \tag{3.24}$$

where n is the number of sub-apertures. We then compute the transformed sub-apertures,

$$\mathbf{S}'_i\{j\} = \mathbf{T}_i\mathbf{S}_i\{j\}, \qquad i = 1 \ldots n, \qquad j = 1 \ldots m. \tag{3.25}$$

where j denotes the j^{th} column and hence point in $\mathbf{S}_i$, and form a single array containing all of the transformed points from each sub-aperture,

$$\mathbf{S}' = \{\mathbf{S}'_i\}, \qquad i = 1 \ldots n. \tag{3.26}$$

Therefore, only when the optimal parameters for stitching have been found on the (much fewer) fiducial points are the full sub-aperture arrays transformed. This represents a considerable saving in time compared to transforming each point in each sub-aperture one or more times per iteration of the least-squares algorithm

The array $\mathbf{S}'$ array contains all the sub-aperture points on n non-aligned equally-spaced grids and represents the stitched measurement result. Points in overlap areas are measured at least twice, and possibly up to 10 or more times depending upon the location of sub-apertures. This implies that the sampling frequency of the data has effectively been increased in the regions of overlap. A technique for dealing with this is developed in chapter 4.

3.3.4 Accessing the Fiducials

Equation 3.23 shows the form of the sub-apertures after the fiducial centroids have been appended, and the positions of these fiducials in the sub-apertures are important in that the distances between matching pairs are evaluated in the least-squares optimization, i.e., the squared distance between fiducials is minimized. A data array format, here named the *fiducial matrix*, has been devised to allow the stitching algorithm to find the coordinates of all pairs of matching fiducials.

Let each fiducial on the physical surface be assigned a number and let there be n_f fiducials in total. We have n_s sub-apertures and therefore can store information

on fiducial pairings in a $n_f \times n_s$ matrix. Referring again to equation 3.23, the final fiducial f_m is stored in the last column of $\mathbf{S}_n$. The column number and sub-aperture number can therefore be used to access any fiducial in any sub-aperture as follows. Let the row index of the $n_f \times n_s$ *fiducial matrix* correspond to the fiducial number and let the column index of the fiducial matrix correspond to the sub-aperture number. The elements of the fiducial matrix, $\mathbf{F}$, are column numbers in the sub-aperture arrays that contain fiducial centroids. As an example, consider figure 3.9 which represents an array of 2×2 square sub-apertures and the associated fiducial points. Fiducial number 1 is in sub-apertures 1 and 3. Therefore in row 1 of the matrix

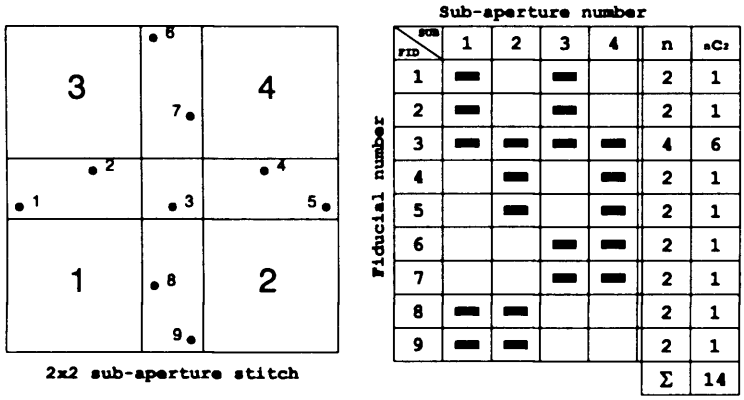


Figure 3.9: Composition of the fiducial matrix. The row index is the fiducial number and the column index is the sub-aperture number. The entries (here black boxes) equal the column number of the fiducial coordinates in sub-aperture n of fiducial m.

there are two entries in columns 1 and 3. $\mathbf{F}_{1,1}$ contains the column index for $\mathbf{S}_1$ that contains the coordinates of fiducial 1. Similarly, $\mathbf{F}_{1,3}$ contains the column index for $\mathbf{S}_3$ that contains the coordinates of fiducial 1 in this particular sub-aperture. Row three contains 4 entries since fiducial number three appears in all four sub-apertures in figure 3.9. By searching along any row, matching fiducials are thus located by the stitching algorithm.

The fiducial matrix can be composed by the centroiding software using the order in which each fiducial was centroided and the machine state parameters as each sub-

aperture was measured. At the time of writing this is unimplemented in software and each fiducial matrix used to produce data in the results section has been composed manually.

3.4 Simulated data generation

After the non-delivery of the OMAM PDWFS in summer 2004 a procedure was developed to generate some simulated surface data that could be used to develop the stitching software, before the work-around using the Fisba and Wyko interferometers was implemented.

A brief and very tentative specification was drawn up for the simulated data. The requirements were as follows:

1. There should be sufficient data such that it could be divided into at least 3×3 sub-apertures, or any other combination of overlapping sub-apertures.

2. It should be generated in a form that is the same as will be read out from the metrology device - that is the simulated sub-apertures should be generated as an 3×n matrices of values in (x, y, z) format.

3. Fiducial points should be picked from the data such that there are sufficient of them for full stitching, i.e., at least 3 per overlap area.

4. The surface should have a 'smooth'-quasi-random form.

The data were generated using a platform-independent C program that simply sends the set of computed points to a text file with each row containing (x, y, z) triplets separated by a space. This text file can be parsed by Matlab and then the transpose can be taken to give the format shown in equation 2.12 on page 119

The program defines several random variables to control the following properties of a random set of 2-D Gaussians:

- The width, in x and y.
- The amplitude.
- The x- and y-position of the peak.
- The total number of Gaussians used.

The program was run and set to produce a regular grid of 200 × 200 points. These points were then output as 16 separate overlapping arrays, each with 70 × 70 points.

For this data, 45 fiducial points were hand picked from the overlap regions to give at least 3 fiducials per overlap area between any two sub-apertures. A clear advantage of having 4×4 sub-apertures is that one can choose any m×n combination of overlapping sub-apertures up to 4×4 without having to completely redefine the fiducials. A coloured image of the simulated surface data, the fiducials and the sub-aperture segments is shown in figure 3.10, where the colour represents the height which can be scaled as desired. The data were used at the development stage of the stitching algorithm to test the procedures developed before using real data.

3.5 Summary

The preceding sections have shown how it is possible to construct a general transformation matrix to apply arbitrary rotation, translation and tip/tilt to a set of points to enable then to be transformed onto another set of matching points in a different orientation or coordinate frame. A summary is now made of the requirements of the stitching algorithm.

1. The data must be pre-scaled, geometrically corrected and in the correct format before stitching using the transformations, $\mathbf{T}_i$, begins.
2. The fiducial matrix must be constructed.

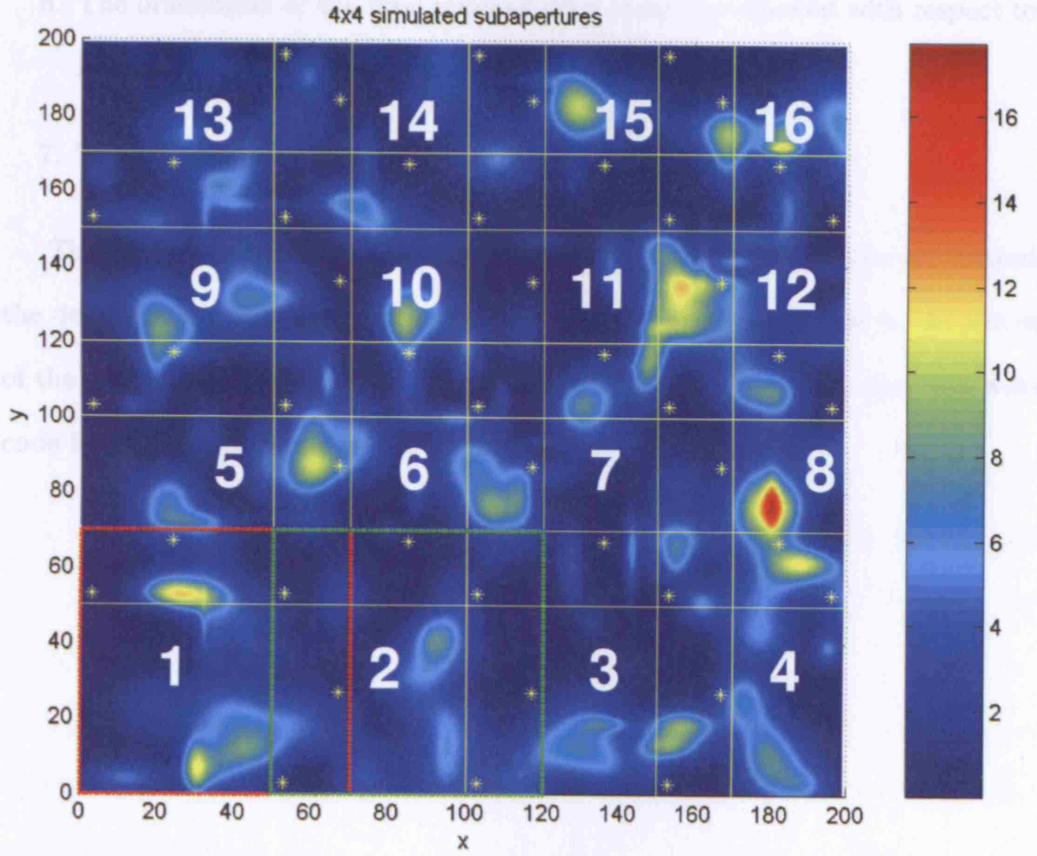


Figure 3.10: Simulated data set as used. Yellow lines denote the sub-aperture edges and yellow stars denote the fiducial positions. The red lines outline sub-aperture 1 and the green lines outline sub-aperture 2.

3. The optimizer must be run to find the optimal stitching parameters.

4. The optimized transform parameters must be applied to the sub-apertures.

5. A single array containing all the transformed sub-aperture arrays must be formed which then constitutes the stitched measurement result.

6. The orientation of the final stitched data must be corrected with respect to a known orientation of the part.

7. The statistics on the surface form must be evaluated.

The next chapter explains the procedures enumerated in 1–7 above and includes the details of the numerical optimization performed to compute the $\mathbf{t}_i$. At the end of the chapter a complete flowchart will be found for the stitching algorithm whose code listings can be found in the appendix.

CHAPTER 4

Stitching Tasks

4.1 Pre-Stitching

This chapter deals with all of the operations to be performed to determine the optimal stitching parameters, from the geometric corrections required to the workings of the optimizer itself.

Much of the previously published work on stitching cited in this thesis does not include such details as scaling data and correcting phasemaps for the particular interferometer transmission element used, for example. There has been a tendency to concentrate on aspects of the techniques used to find the stitching parameters. However, the practical procedures that are ancillary to determining the stitching parameters are equally important and form part of a total process that can be used to pinpoint a particular feature on a measured surface with the required accuracy.

Optical testing is a complicated field and requires a strict methodology when anything other than a flat needs to be tested. Procedures need not only be developed, but followed with care to ensure repeatability and consistency, as evinced by the classic mistakes made during the testing of the Hubble primary mirror [96]. Even

after years of use, some techniques and especially some software solutions have been found to be in error. A particular case came to light at the author's laboratory recently where a scaling factor in a major commercial software package had been defined incorrectly internally to the software, resulting in valleys becoming peaks and vice-versa.

The pre-stitching tasks will be added to the stitching flowchart in the coming sections and code for each stage can be found in the referenced appendices. At the end of the chapter all the necessary steps for stitching will have been explained and justified. The first steps to consider are the geometric corrections to be applied to the sub-aperture data.

4.1.1 Geometric Corrections

Conversion to Cartesian Triplets

All interferometer generated phasemaps are stored as 2-D arrays, as shown in table 2.2 in chapter 2, where the x- and y- coordinates are simply related to the row or column numbers in the arrays. The fiducial centroiding is also done in this coordinate system as per equation 2.11 on page 118. The points must be converted to (x, y, z) form and stored in another array, the $\mathbf{S_i}$. This is achieved by custom Matlab software code that rasters through the 2-D phase array and assigns each point in the new $\mathbf{S_i}$ the value (u, v, ϕ), where u and v are the row and column indices of a point and ϕ is the surface error or form value.

At this time the `NaN` entries in the phasemap are removed such that only valid phase points remain. The fiducial coordinates from the centroiding algorithm are then appended to this array and we now have a sub-aperture in the form of equation 3.23 on page 151.

Aspect Ratio Correction

The Wyko 6000 interferometer uses non-square pixels and when phasemaps are imported into third party software the aspect ratio of the data requires correction. The correct aspect ratio is found by multiplying the y-coordinates by 1.16667. This is accomplished by premultiplying the sub-aperture coordinates in $\mathbf{S}_i$ by a shearing matrix, $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1.16667 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.1}$$

stretching the data into the correct shape. The Fisba interferometer has square pixels and no aspect ratio correction is necessary and the step can be automatically omitted if the Matlab software imports a Fisba phasemap.

Re-centering

Each point in each $\mathbf{S}_i$ is translated by a vector such that the point corresponding to the optical axis in the data is located at the origin - this is only used for spherical parts, but no harm is done on plano parts. This point is found using the alignment mode of the interferometer. Here, a negative lens (shown, for example, in figure 2.8 on page 106) forming part of the interferometer's optical system is shifted into the beam path. This effectively re-images the interferometer's point source onto the CCD. If the SUT's tip/tilt and defocus are nulled, this point is imaged precisely onto the optical axis. Centroiding of the point reveals the coordinates of the optical axis in the frame of the interferometer CCD chip.

Interferometer Image–Object Correction

The phase values from the interferometer are in the u–v plane. Each point, (u, v) has an associated phase/surface value ϕ. We require the coordinates of the object

under test or its error from the reference surface, (x, y, z) from (u, v, ϕ). Let R_{BS} be the radius of curvature of the part's best-fit sphere, and let R_{TS} be the radius of curvature of the interferometer transmission sphere. We have [97]:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{R_{BS}+\phi}{R_{TS}} u \\ \frac{R_{BS}+\phi}{R_{TS}} v \\ R_{TS} - \frac{R_{BS}+\phi}{R_{TS}} \sqrt{R_{TS}^2 - u^2 - v^2} \end{bmatrix}. \tag{4.2}$$

Here, $\frac{R_{BS}+\phi}{R_{TS}}$ is the scale factor between the local measured surface and the reference surface. Thus the u and v coordinates are multiplied by this to give the corrected coordinates. The z-coordinate of the measured surface is found by constructing the z-coordinate of the reference sphere and multiplying this by the local scale factor, $\frac{R_{BS}+\phi}{R_{TS}}$. The z-coordinate of the reference sphere is found from the Cartesian equation of a sphere: $z = \sqrt{R_{TS}^2 - u^2 - v^2}$. The z-coordinate is also shifted along the axis by an amount R_{TS} which can be omitted as this is just a piston term. Furthermore, if we took the positive square root, we would obtain a convex surface instead of a concave one. Figure 4.1 shows the principle of the scaling. Such transformations

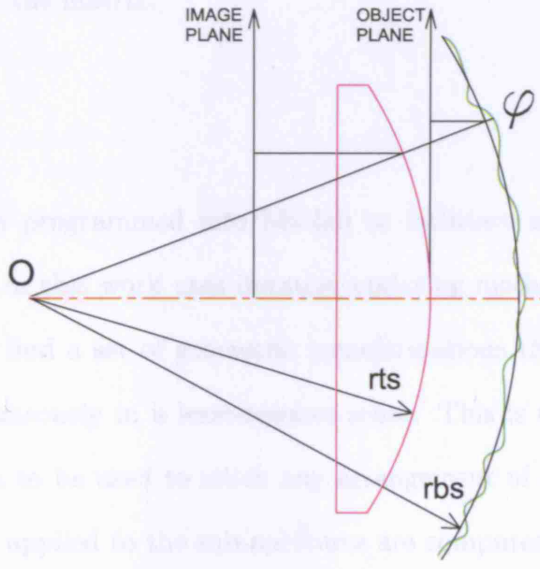


Figure 4.1: Correction of the interferometer phasemap from the part and transmission sphere specifications, redrawn from [97].

are applied to the $\mathbf{S_i}$ to obtain the corrected phasemaps.

Add Homographic Coordinate

The array of corrected sub-aperture values, (x, y, z) must be converted to homographic coordinates, according to section 3.2.1, in order that the transformations in section 3.2.2 can be used. This is accomplished by adding an extra row to the data array. Each entry in this new row has a value of unity in accordance with the mathematics explained in section 3.2.1.

Fiducial Matrix

The fiducial matrix is declared. In an autonomous system, this would be constructed and declared by the software. In this work, the non-availability of a suitable CNC system and time constraints have meant that the fiducial matrix is constructed manually for each sub-aperture testing scheme devised. Thus the order of fiducial centroiding in each sub-aperture must be recorded and used with the sub-aperture map to correctly construct the matrix.

4.2 Algorithms

Two techniques have been programmed into Matlab to facilitate stitching. The main technique developed in this work uses iterative updating methods, described in the coming sections, to find a set of geometric transformations that stitch multiple sub-apertures simultaneously in a least-squares sense. This is envisaged as a generic stitching algorithm to be used to stitch any arrangement of sub-apertures. The transformations to be applied to the sub-apertures are computed as composite matrices consisting of translations, tips, tilts and polar rotations. This algorithm is known as the optimizing algorithm.

Complementary to the main optimization algorithm, a smaller and more re-

stricted algorithm is presented based on a result from linear algebra. The algorithm is useful only for stitching two sub-apertures together at any one time and computes the result directly using the matrix singular value decomposition. The algorithm provides a least-squares stitching between matching fiducial points just like the main iterative algorithm. However, the transformations computed are slightly different in that the tip and tilt operations are approximated by rotations as mentioned in section 3.2.5. This algorithm is known as the Procrustes algorithm. Thus the main differences between the algorithms are in the software method used to find the least-squares fit, iterative versus direct computation, and in the type of transformation found by the algorithms.

The algorithms each have their own merits and problems as shown in table 4.1

Optimizing Algorithm	Procrustes Algorithm
Can stitch any number of sub-apertures per program execution.	Can only stitch two sub-apertures per program execution. Stitching more than two sub-apertures is possible if the algorithm is re-run.
The matrix transformations can be tailored to produce any types of transformation required.	The form of the matrix transformation is fixed.
Consumes relatively large amounts of memory during iteration.	Small memory footprint as only two sub-apertures are considered.
Iterative computation of the transformation matrices takes large amounts of CPU time compared to the Procrustes method – measured in minutes.	Direct computation of the stitching parameters takes very little CPU time – measured in seconds.
Algorithm can be constrained – the upper and lower bounds of the parameters in the matrices can be set by the user.	The algorithm cannot be constrained and under a bad choice of fiducial positions may produce a rotoinversion of a sub-aperture.
Algorithm is more amenable to future upgrades	Algorithm is more difficult to modify

Table 4.1: Advantages and disadvantages of the two algorithms explained in this work.

4.3 Optimization

Now that the pre-stitching tasks have been established, the subject of optimization is considered from a standpoint of producing a practical main algorithm that is simple to implement.

Optimization can be defined quite trivially as *determining the 'best' solution to a mathematically defined problem* [98], according to rules that define the parameters of the best solution. The problem may model something tangible such as a physical system where the first step one must take is to decide what constitutes an acceptable solution to the problem. Then, one can devise an algorithmic method of finding the optimal solution. Murray [99] provides an accessible summary of general methods of numerical optimization but much of the content in this chapter draws from the information in the *Matlab Optimization Toolbox users guide* [77].

4.3.1 Terminology

A function to be optimized is called the *objective function* and the usual convention is to find the minimum value of this function. The object function may contain several – or even hundreds –of variables and many of these variables may be constrained. The general problem of optimization is expressed thus:

> Given a function $f(\mathbf{x})$, where $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, find the values of $\mathbf{x}$ that make $f(\mathbf{x})$ a minimum, and denote this point by $\overset{*}{\mathbf{x}}$.

It is easy to see that minimization and maximization are related very trivially: one could minimize f or maximize $-f$ therefore this kind of optimization is usually referred to as *numerical minimization*. It is highly desirable that the minimization takes as little time as possible and uses the least amount of computer memory possible.

The simplest case is for a univariate function, where the function, say $f(x)$, is only defined over an interval $[a, b]$ and the stationary points are found by solving $\frac{df(x)}{dx} = 0$

if they exist over the interval. Several types of extrema may be found for a given function and figure 4.2 illustrates the important types that may be encountered. In

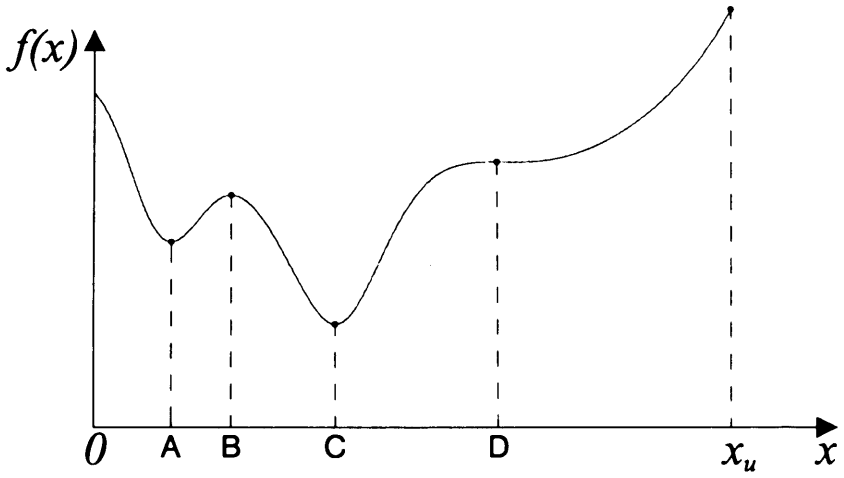


Figure 4.2: Some important extrema of a function of a single variable. The global minimum is at point C, and the global maximum at point x_u

figure 4.2, $f(x)$ is defined $\forall x \in [0, x_u]$. Point A represents a *strong local minimum*, point B represents a *strong local maximum*, and point C is the *global minimum*. Point D is a point of inflexion and since $f(x)$ is defined only on $[0, x_u]$, the point at x_u is the global maximum. A minimum may therefore be the true global minimum or a minimum over a subset of the domain of $f(x)$, such as that at A in $[0, B]$ in the figure. When more complex functions are considered it is not always possible to determine if a minimum is the global minimum.

There are two simple approaches [100] to the problem of optimization that have resulted in a variety of studies in the general field of numerical research. The first involves choosing varying starting values in a random or quasi-random manner and picking the most extreme value produced. The second involves starting at a point close to an extremum, then choosing a point near to this and checking if the second point results in a lower value of the object function. In general the function to be optimized is a function of many variables and the optimization process is equivalent to finding the lowest point on a surface described by the object func-

tion. The procedure is relatively straightforward to describe: pick a start point and travel over the surface following the direction of *steepest descent* until a minimum is found. This technique can be quite slow and modern techniques generally tend to use a combination of methods. For example, the well-known Levenberg-Marquardt method uses a search direction that the identical to the steepest descent far away from the minimum and uses a Gauss-Newton (chi-squared) algorithm near to the minimum [77].

It is often more difficult to find the global minimum than it is to locate a local minimum. One way to check is to vary the starting parameters to determine if there exists several minima and compare the values of the minima.

Figure 4.3 shows an example of a 2-D surface, known as Rosenbrock's function and its global minimum $\forall x, y \in \mathbb{R}$. The minimum is located at $(1, 1)$. Modern computer maths software has come a long way since the early days of stored-program computers and many commercial packages have general algorithms for optimization that can be applied easily once the problem has been posed in a suitable form. As an example, the calculations and plot shown in figure 4.3 were programmed in just 8 lines of user code. The Matlab functions called by the user code, in contrast, were much longer requiring many lines of code, and implemented a Newton-type search by numerical evaluation of the gradients along the direction of search.

4.3.2 Notation Used

Since many optimization methods are iterative it is useful to define a few terms as used in the next sections. Let $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ be a set of points $\{\mathbf{x}^{(k)}\}$, where k is the iteration number. It is desirable that these points converge on the point denoted by $\overset{*}{\mathbf{x}}$, which is a solution to the problem. In later sections, the importance of a line with respect to a line search will be dealt with. The set of points, $\mathbf{x}(\alpha) = \mathbf{x}^{(k)} + \alpha\mathbf{s}$, defines a line (for all $\alpha \in \mathbb{R}$) where $x^{(k)}$ is a fixed point, α is the step and $\mathbf{s}$ is the

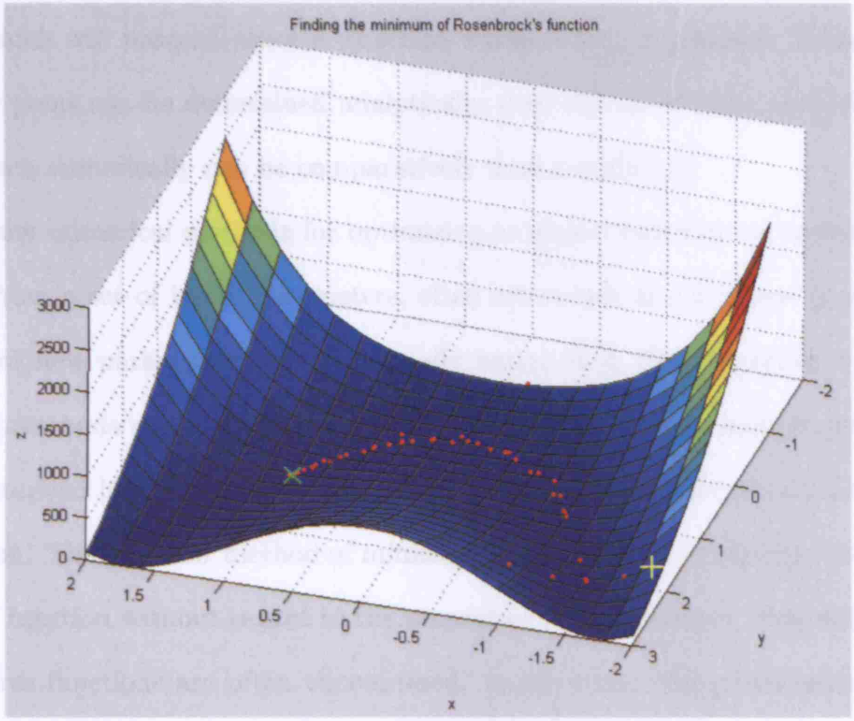


Figure 4.3: Optimization of Rosenbrock's function; start point: yellow plus, evaluation points: red dots; end point: green x at $(1, 1)$.

direction of the line. We have a half line for all $\alpha \geq 0$.

Further constructions of use when $f(\mathbf{x})$ has continuous first and second derivatives are the *gradient vector* and the *Hessian*. If ∇ denotes the gradient operator $(\partial/\partial x_1, \cdots, \partial/\partial x_n)^T$, then $\nabla f(\mathbf{x})$ is the gradient vector, and $\nabla(\nabla f^T) = \nabla^2 f(\mathbf{x})$ is the Hessian, sometimes denoted by $\mathbf{G}$ and $\mathbf{H}$ respectively. The gradient and Hessian are important constructions in optimization problems, especially if they can be computed analytically for $f(\mathbf{x})$. For example, in a steepest descent search, the line search will proceed along a direction where $\nabla f(\mathbf{x})$ is greatest. If the gradient at any point can be determined analytically then this saves time, since finding the gradients numerically can be comparatively time consuming.

Many numerical methods for optimizing an object function are iterative. They start from a set of initial parameters, often estimated, and progress by generating a set of new parameters that are usually better than the preceeding set. *Direct search* methods are often amongst the simplest types of optimization and they are characterized by the fact that they do not make use of any derivatives of the object function. That is, their method of minimization is based on comparing values of the object function without regard to the properties of its derivative. Non-differentiable objective functions are often encountered, as are those with a discontinuous first derivatives.

4.4 Optimizing Algorithm: The Object Function

The focus of the present section is to define the properties of the object function for the optimization in terms of the data and the transformations to be applied.

Section 3.2.2 established the form of the transformations that are required to stitch the sub-apertures. The sum of the squared distances between fiducial pairs will be minimized after these transformations are applied and the data are said to be stitched after this optimization. A very comprehensive toolbox is available for

Matlab for solving an optimization problems such as this. These tools, from the *Optimization toolbox*, have been used in this work since it is generally quicker and less troublesome, and therefore practical, to use a ready-made algorithm than to try to embody all of the current knowledge on a subject into a program of one's own. Most of the stitching optimizations in later sections are performed by Matlab's *Optimization toolbox* functions `fmincon.m` and `fminunc.m` for constrained and unconstrained minimizations respectively. The functions are operationally identical except for the fact that `fmincon` accepts extra arguments that specify bounds for the design variables.

The idea presented in this chapter is to compute and apply a transformation matrix to be applied to each sub-aperture. Following this, the object function is evaluated which simply sums the squared distances between fiducial pairs as a function of the transformation parameters. The algorithm then attempts to compute a better set of transformations based on the gradient of the object function at the current point, $\mathbf{x}^{(k)}$, and attempts to lower the value of the object function. When the object function is low enough or the change in the value of the object function is low enough, the optimizer terminates.

In the past, a computation such as this would have required a considerable investment in computing power to compute and apply all the transformation matrices and the corresponding gradients. Today with high-speed desktop computing and ample memory, brute force methods such as this can be applied more readily.

Equation 2.8 on page 90 describes the stitching equation to be minimized for only the fiducial points between two sub-apertures. Here, each time the object function is evaluated, pairs of fiducials are found and the distances between the partners in a pair are squared and added over all the sub-apertures simultaneously. This leads to the procedure for forming and evaluating the object function:

Guess: Form the initial guess for the transformation parameters $\mathbf{t}_i$ used to seed the

optimizer.

Transforms: Compute the first or subsequent transformation matrices $\mathbf{T_i}^{(k)}$, defined in equation 3.18, for each sub-aperture and based on the parameters $\mathbf{t_i}^{(k)}$.

Transform: Transform only the fiducial points in each sub-aperture by the appropriate transformation matrix for that aperture.

Residual: Calculate the total fiducial residual, f, which is defined as the sum of the distances squared between all fiducials in each pair, for all pairs.

4.4.1 Guess

The initial value, or guess is where the optimizer begins in its search for the minimum. In practice, for an unconstrained optimization, the initial guess could be totally random, but this may mean that the chance of finding a local minimum is increased. Furthermore, if the initial guess is close to the values required to find the minimum then the optimizer will terminate more quickly as less iterations will be required. For a constrained optimization, the initial guess must be within the bounds supplied to the optimizer function. For this work, the initial guess values were based on the configuration of the interferometer and the surface under test at the time of each measurement. In a CNC system the machine's position encoders would simply be read out each time a sub-aperture was measured and fed into the stitching algorithm.

4.4.2 Transforms

The second part of the object function contains steps necessary to form the transformation matrices for each sub-aperture. The parameters for each transformation change at each iteration and therefore new matrices are computed at each step with

the new parameters for the k^{th} iteration provided by the previous iteration of the optimizer.

The trigonometric functions in each $\mathbf{T_i}$ are evaluated only once per matrix per iteration to give a 4×4 matrix of reals.

4.4.3 Transform

In step 3 the fiducial coordinates are simply pre-multiplied by the appropriate transformation matrix to give their new positions. The fiducial coordinates are found in the sub-aperture data by reference to the *fiducial matrix* as described in section 3.3.4. The optimizer maintains the new coordinates of the fiducials in local variables only.

4.4.4 Residual

A nested loop is executed that locates the transformed fiducial pairs and computes and sums the squared l^2 norms of the vectors defining the distance between fiducials in pairs. This sum is returned by the object function to the optimizer to be compared with the value for the next iteration. When the sum is deemed small enough the optimizer terminates.

4.4.5 Calculation of the Fiducial Residual

Each fiducial appears in at least two sub-apertures and let us call the same physical fiducial in the two different sub-apertures a pair. A fiducial may therefore be a part of more than one pair. After a transformation the 3-D distance remaining between a pair of fiducials is computed as the l^2 norm of the vector joining them. The norm is then squared to give a positive value and this is then added to the total sum including all distances between paired fiducials. This sum is computed by a 3-deep nested loop with two levels of decisions, to parse the fiducial matrix via some very daunting indexing as revealed in appendix G.

The exact number of fiducial pairs to consider can be computed directly from the fiducial matrix as follows. For a fiducial appearing in n sub-apertures, the number of distances to square and add is given by the binomial coefficient

$$ {}_nC_2 = \left(\begin{array}{c} n \\ 2 \end{array} \right) = \frac{n!}{2!(n-2)!} = n_p, \tag{4.3} $$

since fiducials are evaluated pairwise. This knowledge of the number of pairings per fiducial allows the total number of sums per optimizer iteration to be evaluated which in turn allows the mean distance between fiducial pairs to be computed. We have the object function value, f

$$ f = \sum_{\text{pairs}} \text{distances}^2, \tag{4.4} $$

therefore,

$$ (\text{mean distance})^2 = \frac{f}{n_p}, \tag{4.5} $$

and therefore, $\bar{r}$, the mean distance between pairs of fiducials is given by,

$$ \bar{r} = \sqrt{\frac{f}{n_p}} = \text{mean distance}, \tag{4.6} $$

where n_p is the number of pairs considered in the sum and f is the object function value at the current iteration point $\mathbf{x}^{(k)}$. n_p is computed from the fiducial matrix by finding the number of sub-apertures in which each fiducial appears and by finding how many distinct pairs can be made (${}_nC_2$) for each row (i.e. fiducial). The sum of all the combinations gives the total number of pairs to consider as shown in figure 3.9 on page 155.

The numerical methods used here can be used to set the parameters for a successful optimization termination. After a successful stitching run has been made, equation 4.6 can be used to compute a value for f for the new stitching scheme where n_p is known from the arrangement of sub-apertures and fiducials. Thus

$$ f = n_p \bar{r}^2 \tag{4.7} $$

gives a target f-value for a new stitching scheme (i.e. a different part) as derived from the successful stitching of a previous part under similar circumstances.

The ideas presented here can also be used to estimate the likely f-values encountered by consideration of the random and systematic errors of the measurement and fiducial centroiding processes, and examples of these can be found in chapter 6.

4.5 Optimizing Algorithm: The Process

The Matlab *Optimization toolbox* functions `fminunc.m` and `fmincon.m` seek to find the minimum of an unconstrained or constrained multivariate function:

$$\min_{\mathbf{x}} f(\mathbf{x}), \tag{4.8}$$

where $\mathbf{x}$ is a vector and $f(\mathbf{x})$ returns a scalar, f, which is the object function value. The optimizer needs access to the object function code, the fiducial matrix, the sub-aperture data maps, the initial guess for the transform parameters, bounds for $\mathbf{x}$ if constrained optimization is used and the options concerning the preferred type of optimization.

Object Function: The object function is a procedure (code) to be used by the optimizer to find the object function value at each iteration for a given set of input parameters.

Fiducial Matrix: A document describing the pairings of fiducials and the locations of their coordinates in the sub-aperture arrays. Figure 3.9 on page 155.

Sub-aperture data: The measurement patches containing the surface data points and the fiducial points, $\mathbf{S}_i$.

Seed: The initial parameters, or start point for the optimizer, $\mathbf{t}_i^{(0)}$.

Bounds These are simply two vectors containing the lower and upper bounds for the design variables. These are of the form $[A_l, B_l, \psi_l, x_l, y_l, z_l]^T$ and

$[A_u, B_u, \psi_u, x_u, y_u, z_u]^T$.

Options: The configuration options for the optimizer subroutine.

The options passed to the optimizer control the exact algorithm and the parameters used to determine when to stop the optimizer. In this work the following options are used with varying values:

Function tolerance: The optimizer terminates if the change in the function value over 1 iteration is less than the value specified.

x-value tolerance: The optimizer terminates if the change in the value of x over 1 iteration is smaller than the value specified.

Finite differencing max and min: This option controls the maximum and minimum step sizes used in determining the directional derivatives.

Max. function evaluations: The optimizer will terminate if the number of object function evaluations reaches this value.

Max. iterations: The optimizer will terminate if the number of iterations reaches this value.

In general the number of function evaluations per iteration is greater than 1 since numerically derived gradients are computed during each iteration. The finite difference method involves perturbing each of the design variables in turn and calculating the rate of change of the object function. Supplying analytical gradients can be beneficial to the optimizer since a direct analytical computation of the gradient at **x** is more efficient than finite differencing. However, the form of the object function does not allow the computation of analytical gradients in this instance.

The particular routines use a quasi-Newton method employing a mixed quadratic and cubic line search method. This simply means that when the line search algorithm is operating, points on the chi-squared surface are extrapolated or interpolated using

quadratic or cubic methods as amply explained in [77]. The routine is constrained to operate on functions that are continuous and real which does not pose any problem for the object function since the return value is a real scalar formed by a sum of real distances in $\mathbb{R}^3$.

The procedures, `fminunc` and `fmincon`, build up curvature information at each iteration to form a quadratic model of the problem as stated in [101]:

$$\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{c}^T\mathbf{x} + b. \tag{4.9}$$

The optimum solution occurs when the partial derivatives of $\mathbf{x}$ equal zero

$$\nabla f(\overset{*}{x}) = \mathbf{H}\overset{*}{\mathbf{x}} + \mathbf{c} = 0, \tag{4.10}$$

where the optimum solution point $\overset{*}{x}$ is therefore equal to

$$\overset{*}{x} = -\mathbf{H}^{-1}\mathbf{c}. \tag{4.11}$$

Newton methods compute $\mathbf{H}$ directly, but the numerical calulation of $\mathbf{H}$ can be very time consuming. The Quasi-Newton method avoids long computations by using the observed behaviour of $f(\mathbf{x})$ and $\nabla f(\mathbf{x})$ to successively approximate $\mathbf{H}$ using an updating technique.

At each major iteration a line search is then performed in the direction

$$\mathbf{d}_{(k)} = -\mathbf{H}^{-1} \cdot \nabla f(\mathbf{x}^{(k)}). \tag{4.12}$$

The method finds

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)}, \tag{4.13}$$

where $\mathbf{x}^{(k+1)}$ is the next iterate and α is a scalar step length parameter. The line search thus attempts to decrease the objective function along the line $\mathbf{x}^{(k)} + \alpha\mathbf{d}$.

4.5.1 The software

Let $\mathbf{F}_i$ be a subset of $\mathbf{S}_i$ but containing only the fiducial points that were previously appended to the end of the $\mathbf{S}_i$. The fiducials are transformed:

$$\mathbf{F}_i^{(k)} = \mathbf{T}_i^{(k)}\mathbf{F}_i, \tag{4.14}$$

, where $^{(k)}$ denotes the k$^{\text{th}}$ iteration. The points have the form:

$$\begin{vmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \\ z_1 & \dots & z_n \\ 1 & \dots & 1 \end{vmatrix} \rightarrow \begin{vmatrix} x'_1 & \dots & x'_n \\ y'_1 & \dots & y'_n \\ z'_1 & \dots & z'_n \\ 1 & \dots & 1 \end{vmatrix}. \tag{4.15}$$

Let the m$^{\text{th}}$ fiducial in $\mathbf{S}_i$ also be located in $\mathbf{S}_j$ as the n$^{\text{th}}$ fiducial. The squared fiducial residual for the pairing is calculated as

$$f_{i,j,m,n} = \left| \begin{pmatrix} x_m \\ y_m \\ z_m \end{pmatrix}_i - \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}_j \right|^2 \tag{4.16}$$

The cumulative sum of all pairings is then computed as the scalar value for the object function:

$$\sum_{\text{pairs}} \left| \begin{pmatrix} x_m \\ y_m \\ z_m \end{pmatrix}_i - \begin{pmatrix} x_m \\ y_n \\ z_n \end{pmatrix}_j \right|^2. \tag{4.17}$$

With ideal data it should be possible for f to evaluate to zero if the transformations are exact and no rounding errors occur in the computer. Real data will of course involve positive values of f whose magnitude will be determined by

1. The number of fiducial pairs to sum, i.e, the number of sub-apertures to stitch.
2. The errors in fiducial centroiding.
3. The random errors produced from noise sources (turbulence, vibration and stray light).
4. The termination criteria for the optimizer.
5. Rounding errors in the computer.

The termination conditions for the optimizer are set by the `options` input and whenever the optimizer terminates the following are output:

$\mathbf{t}_i^{(\mathrm{opt})}$: The values of the stitching parameters at the time of optimizer termination used to construct the optimal stitching transformations $\mathbf{T}_i^{(\mathrm{opt})}$.

f: The value of the object function when the optimizer terminated.

EXITFLAG: The result of the optimization (i.e. success or failure reason) in the form of a numerical value.

output: A structure containing the optimization history and process statistics.

Figure 4.4 illustrates the inputs and outputs to the optimizer.

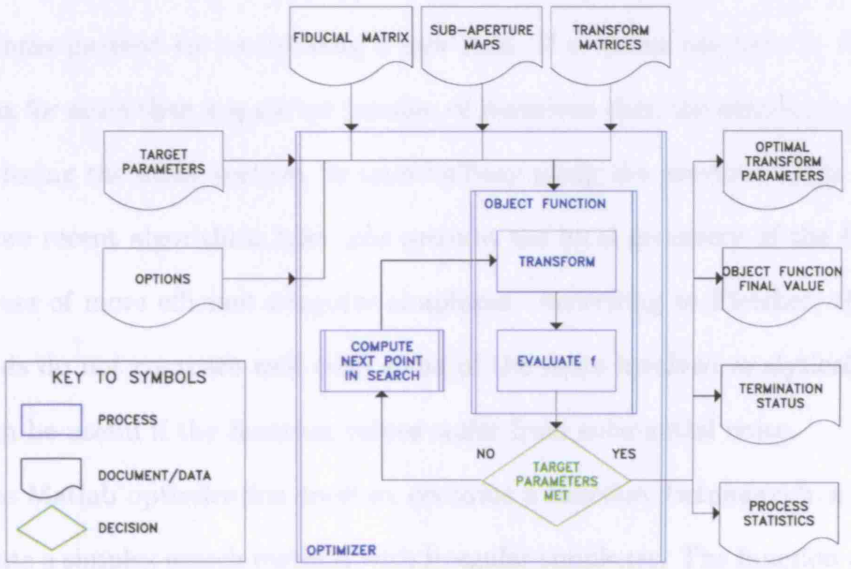


Figure 4.4: Diagram showing the inputs and outputs to the optimizer. `Fminunc` is represented by the large central box labelled 'OPTIMIZER' and the object function is represented by the nested box, 'OBJECT FUNCTION'.

4.5.2 A Quick, Dirty and Free Alternative

Fletcher [98] mentions so-called *ad hoc* methods for minimization, based on simple concepts with limited theoretical background. One such method merely involves

picking random points in a fixed region and selecting the optimal point after a large number of trials. For simple low-dimensional problems, bisection methods can be used, where an attempt is made to contract the region systematically. These methods can be extended to n-dimensions and the most successful is known as the *simplex method* first suggested by Spendley and described by Fletcher [loc cit.].

A simplex is a set of n+1 equidistant points in $\mathbb{R}^n$ and at the commencement of optimization, the object function is evaluated at each point or vertex. The vertex where the function value is the largest is reflected in the centroid of the other vertices to form another simplex. Then the process is repeated. If the newest vertex still has the largest function value then the vertex with the next largest value is reflected. Ultimately, a point will be reached where no progress can be made and so the algorithms proceed by introducing a new rule. If a vertex has been in the current simplex for more than a specified number of iterations then the simplex is contracted by replacing the other vertices by ones halfway along the previous edges.

More recent algorithms take into account the local geometry of the function to make use of more efficient irregular simplexes. According to Fletcher, the simplex methods do not compare well with some of the more involved analytical methods, but can be useful if the function values suffer from substantial noise.

The Matlab optimization toolbox contains a function `fminsearch.m` which implements a simplex search method with irregular simplexes. The function arguments and object function can all take the same forms as `fminunc.m` and `fmincon.m` and therefore they can be replaced by `fminsearch.m` without any software modification except for the removal of the constraints if `fmincon.m` was previously used.

4.6 Procrustes

Equation 2.8 on page 90 is the statement of a minimization problem often referred to as a *Procrustes problem*. Procrustes [60, 102] is a notorious character from Greek

mythology who kept an inn at Attica with a unique bed for his guests. He claimed that the bed would fit whomsoever lay down upon it. Unfortunately for Procrustes' guests, they were made to fit the bed perfectly – tall 'guests' had their limbs lopped and short 'guests' were lengthened by racking. Variations of the myth allow Procrustes two beds or one bed that is adjustable - this is to exclude Procrustes from the social embarrassment of having an initially perfect-fitting guest. Procrustes was finally fitted to his own bed by Theseus in a crackdown on robbers and bandits.

Gower and Dijksterhuis provide an excellent overview of Procrustes methods and their opening paragraph serves to illustrate the story in a humorous, albeit algebraic way [60]:

> "There are three elements in the Procrustes story, the unfortunate traveller, who we might label $\mathbf{X}_1$, the Procrustean bed $\mathbf{X}_2$, and the treatment, $\mathbf{T}$, meted out of racking, hammering, or amputation. The simplest algebraic treatment of a Procrustes problem seeks a matrix $\mathbf{T}$ that minimises
>
> $$\|\mathbf{X}_1\mathbf{T} - \mathbf{X}_2\| \tag{4.18}$$
>
> over $\mathbf{T}$ for a given $\mathbf{X}_1$ and $\mathbf{X}_2$.", from Gower and Dijksterhuis, *Procrustes Problems*.

This can be restated as determining a linear transformation $\mathbf{T}$ of the points in $\mathbf{X}_2$ to conform them to the points in $\mathbf{X}_1$ with the least mismatch.

In general, Procrustes problems can conform to the ideas of Procrustes himself – brutal methods can be used to make sets of points in an n-dimensional space fit another set of points. The transformation $\mathbf{T}$ in equation 4.18 can however be made orthogonal such that deformation of the points does not occur.

Procrustes analysis normally matches the i^{th} point in $\mathbf{X}_2$ to the i^{th} point in $\mathbf{X}_1$. The Matlab statistics toolbox contains a function *procrustes.m* that can be used to

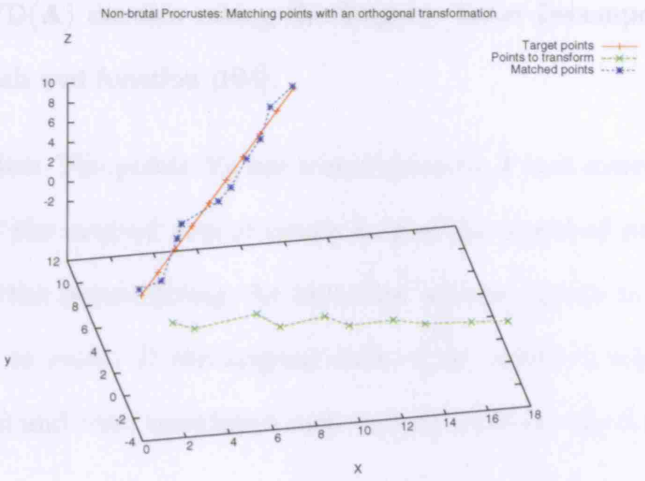


Figure 4.5: Bed fitting the Procrustean way. The green set of points is a rotated copy of the red set with added noise. The green points are then treated to a Procrustean fitting to give the matched (in a least-squares sense) points in blue. In the absence of noise, the fit is limited by the precision of the computer's float storage precision.

determine the transformation between two sets of points. This can be useful when only two sub-apertures of nulled measurements are to be stitched where the solution is directly computed by a matrix decomposition method. The idea can be extended to $n \times 1$ sub-apertures. The theory behind the transformation determination is quite involved, see for example Seber [103] or Gower and Dijksterhuis [60], but the steps can be summarized as follows. Let $\mathbf{X}$ be the target points, and $\mathbf{Y}$ be the points to match to $\mathbf{X}$.

Scale and center Each set of points is centred about the origin and scaled such that it has a unit norm to give $\mathbf{X}_0$ and $\mathbf{Y}_0$

Check dimensionality If $\mathbf{Y}$ has less points than $\mathbf{X}$ then pad with zeroes.

Optimum transformation The optimum transformation matrix is given by

$$\mathbf{A} = \mathbf{X}_0' \mathbf{Y}_0 \tag{4.19}$$

$$[\mathbf{U} \quad \mathbf{S} \quad \mathbf{V}] = \mathrm{SVD}(\mathbf{A}) \tag{4.20}$$

$$\mathbf{T} = \mathbf{V}\mathbf{U}. \tag{4.21}$$

Here, SVD(**A**) denotes taking the Singular Value Decomposition of **A** using the Matlab `svd` function [104].

Transformation The points $\mathbf{Y}_0$ are transformed by **T** and scaled according to the norms of the original sets of points to give the matched points. The scaling factor of the points giving the best least-squares match to **Y** may not quite be equal to unity. If the original scale of the points is required then **TY** is computed and then translated such that its centroid matches that of **X**

Figure 4.5 shows an example of two sets of points that were matched using the above procedure.

This procedure can be applied to the known fiducials between two sub-apertures to provide a non-iterative and therefore quick solution to the stitching of any two sub-apertures. The principle can be extended to measure any $n \times 1$ sub-apertures by sequential stitching of the next sub-aperture to the already stitched sets. This could be used to stitch measurements on long flats or calibration bars, for example. However, the algorithm needs to be used with some degree of caution with regard to the tip and tilt terms in the data.

The Matlab routines programmed involve the same pre-stitching steps as before, including the fiducial centroiding. It is to be noted, however, that the algorithms proceed via rigid rotation matrices and can be restricted to stitching where the tip and tilt angles are small.

4.7 Post Stitching

The section is devoted to the procedures necessary once the stitching transformations have been derived. The processes include overall tip/tilt correction, alignment (orientation), filtering and analysis.

The sub-apertures are first transformed by the appropriate transformation ma-

trix:

$$\mathbf{S}_i' = \mathbf{T}_i\mathbf{S}_i. \tag{4.22}$$

Now, the set of sub-apertures $\{\mathbf{S}_i'\}$ is merged into a single array which we shall call $\mathbf{S}$. The surface form or form error in $\mathbf{S}$ will, in general, be in a random orientation compared to how the part was measured. The orientation can be corrected by fitting a plane to the data and removing residual tip/tilt errors for the whole surface, followed by a final rotation about z to align the part for convenience and for plotting the error maps.

4.7.1 Orientation Correction and Centering

The orientation of the stitched data can be restored by fitting a plane to the data and establishing the parameters required to orient it in a meaningful way - usually parallel with the x- and y-axes. Singular Value Decomposition (SVD) again provides a convenient method of achieving this. A plane is fitted to the data in a least-squares sense, that is, the sum of the distances of all data points from the plane is minimized.

The singular value decomposition of the data is taken using the Matlab function `svd.m`. The function returns three matrices, one of which contains the basis vectors for the space spanned by the data about the best-fit plane defining the residual tip and tilt [104]. Following this correction, the data are re-centered about the origin. The procedure is as follows:

1. Compute the singular value decomposition of $\mathbf{S}$. This returns three matrices:
   $[\mathbf{u}\ \mathbf{s}\ \mathbf{v}] = \mathrm{SVD}(\mathbf{S})$
2. The matrix $\mathbf{u}$ contains the basis vectors of the space spanned by the data in the form

$$\mathbf{u} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ y_{21} & \ddots & \vdots \\ z_{31} & z_{23} & z_{33} \end{pmatrix}, \tag{4.23}$$

here, the third column contains the vector normal to the best-fit plane which we shall call $\mathbf{n} = (n_x, n_y, n_z)^T$.

3. The direction of $\mathbf{n}$ gives the gradients in the x-z and y-z planes.

4. The data is tip/tilt corrected by removing this plane from the data.

5. The polar rotation is corrected to orient the data 'the right way up' by applying a rotation matrix:

$$\mathbf{S}_{oriented} = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \cdot \mathbf{S}. \tag{4.24}$$

6. The data centroid is evaluated in x, y and z, that is, the means of the phase values in x, y and z are found. The data are then translated along this vector to give the origin at the part centre. Let $\mathbf{c} = (x_c, y_c, z_c)^T$ be the centroid. The data are ready for final analysis after applying

$$\mathbf{S}_S = \mathbf{S}_{oriented} - \mathbf{c}. \tag{4.25}$$

4.7.2 Regridding

Prior to filtering, input into the form control software of the Zeeko polishing machines or to plotting, the data must be regridded. Any two overlapping sub-aperture maps will contain a higher density of points in the overlap area on two non-coincident grids. The regridding resamples the data onto a uniform grid and must ensure that the resulting regridded data passes through the original points.

Regridding is performed by the Matlab function `griddata.m` and results in three 2-dimensional arrays, one each for x, y and z. The grid spacing is chosen by the user to ensure that the original data are sampled at a rate appropriate to the original sub-aperture measurements. This can be found from the interferometers (u, v) resolution unit of the part at measurement time.

Up to this point in the processing, there have been 'holes' in each sub-aperture data set corresponding to the fiducials. The surface data are unknown at these points, but the regions are very small compared to the size of the whole surface. However, if a measurement is required at these points, then two stitching measurements will have to be made, with the second set using a different configuration of fiducials such that none are located in positions occupied in the first set. The Matlab function `griddata.m` interpolates the surface at the fiducial points using a linear model based on the shape of the surface surrounding the data hole.

4.8 Filtering

Filtering forms an important part of the analysis procedure to determine a part's form and produce an acceptable measurement result. Surface measurement can be broadly divided into texture and form measurement. The distinction between the two simply concerns the spatial frequencies involved and the question arises regarding the value of the transition frequency from texture to form. This depends on the particular properties of the surface required and how the international standards are interpreted [105]. *Texture or roughness* is a measure of how rough a surface is and may encompass a range of spatial frequencies from 0.1 mm^{-1} up to infinity or the upper limit of the measurement device. *Form* is a measure of the lower spatial frequencies - that is the frequency components of the surface that dictate its overall shape from some upper frequency limit down to zero. The distinction also applies to the magnitudes of the frequency components in the power spectral density. Texture magnitudes are generally smaller than the form magnitudes unless very rough ground surfaces are considered.

The stitching process inevitably introduces spatial frequencies to the data higher than those that the interferometer can measure for the simple reason that when two sub-apertures are stitched in the presence of noise and a translation between sub-

apertures that does not equate to an integer number of pixels on the interferometer CCD, there will be two equally spaced grids of data that overlap.

The discrepancy between two phasemaps of the same part using the interferometer has been found to be lie in the range 5–10 nm RMS, see figure 4.7, and from examination of the plots in the results chapter. Thus, when two grids of data are combined, there will be a spatial frequency component with a spacing (wavelength) of less than one half times the resolution imprinted on the surface under test. The magnitude of the discrepancy between two overlaps can be used as an estimate of the residual stitching errors, as used by Fleig et al in the QED stitching interferometer. In this work, the fiducial residual, or object function value at the end of the optimization is used as a metric of stitching error.

The spatial frequency component from overlapping data has been found in stitched data sets and removing it requires filtering. Figure 4.6 shows a patch of data from a stitched measurement result. The black lines indicate where sections have been taken through the data in x and y to produce the plots in figure 4.7. The plot clearly shows the effects of combining the two sets of data. Figure 4.7 on page 189 shows the plots of the sections in x and in y and the Power Spectral Density (PSD). The PSD plots clearly show the stitching spatial frequencies at approx 48 per thousand lateral units and harmonics thereof.

Many types of filters can be used but for this work, the simplest to implement were found to be simple convolution filters. The most basic filter of this type weights each pixel by the neighbouring pixels. This could mean, for example, that a pixel's new value after filtering is equal to the mean of the surrounding 1, 2 or n layers of pixels. The weighting *kernel* is convolved with the data points to produce the filtered data.

A kernel with a Gaussian profile was computed and applied to the data in figure 4.6 to give the filtered patch shown in figure 4.8. The filter kernel used was a

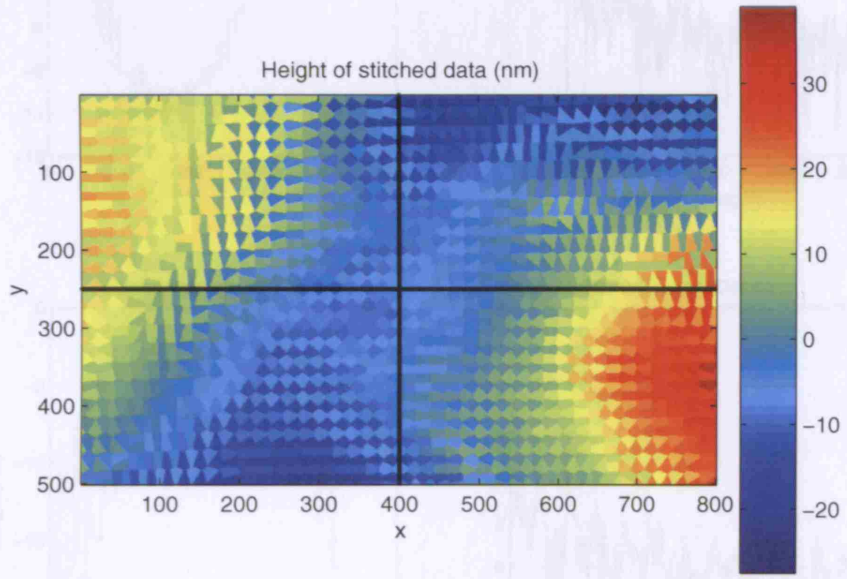


Figure 4.6: Stitched data patch showing the effects of combining two data sets with interspersed grid points. The net effect is to generate high spatial frequencies in the data.

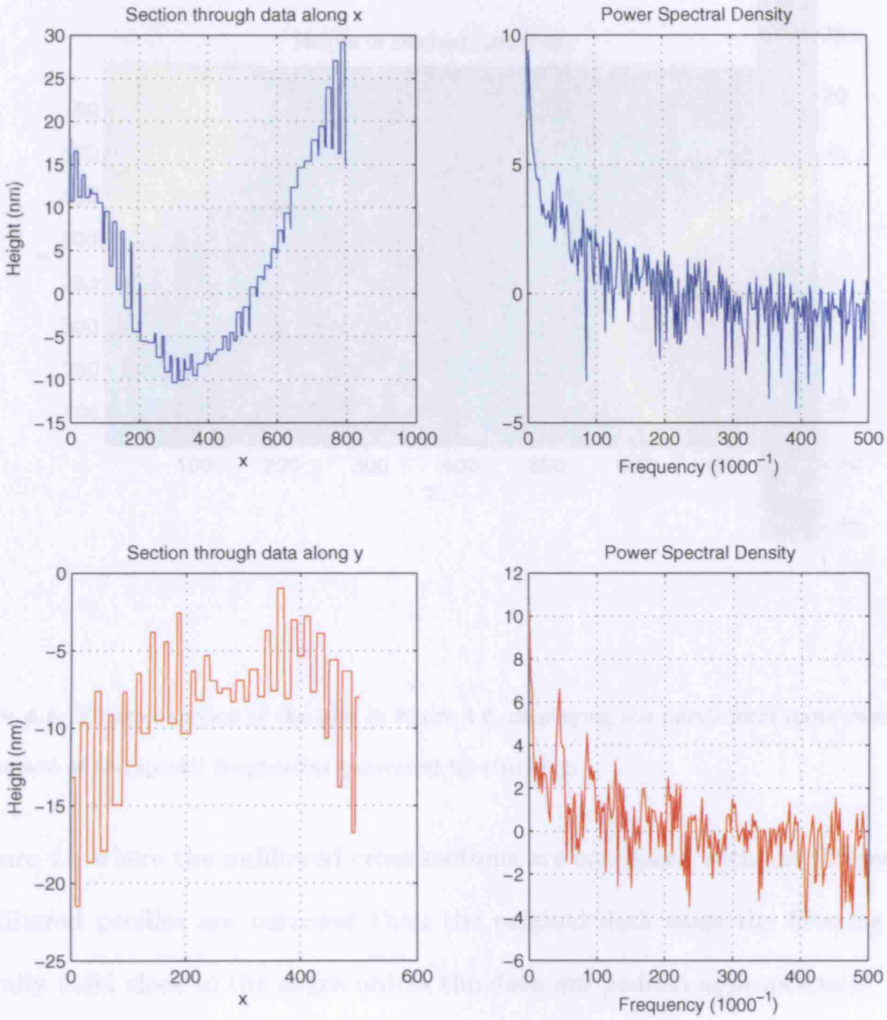


Figure 4.7: Cross section plots thorough the data shown in figure 4.6 on page 188 and the corresponding PSD. The PSD clearly shows the high spatial frequencies generated by the stitching at approximately 48 per 1000 units and harmonics of this. For the purposes of this measurement, form frequencies extend from DC up to 10 per 1000 units.

normalized 64×64 matrix of points of a Gaussian form with a FWHM of 33 pixels. Post filtering, the form of the patch is more apparent. This is more clearly shown

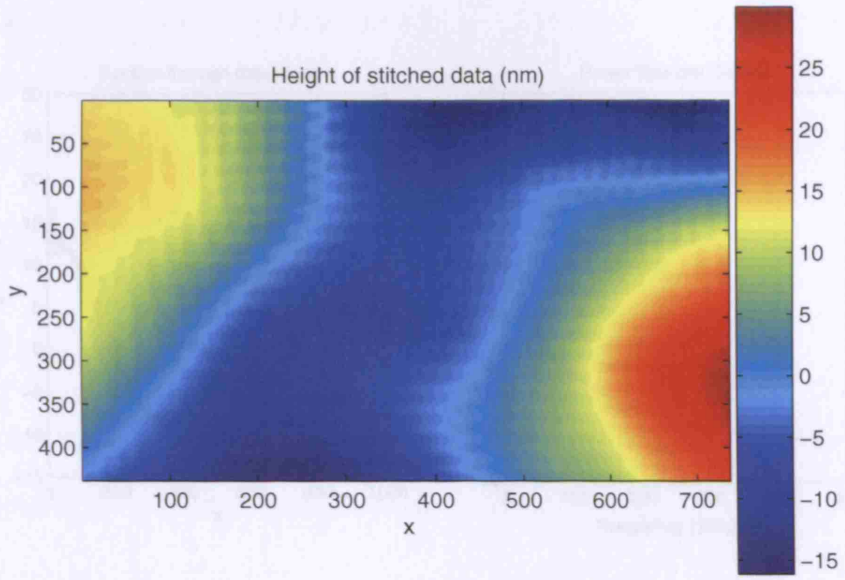


Figure 4.8: Filtered version of the plot in figure 4.6, displaying the part's form more clearly after the removal of the spatial frequencies generated by stitching.

in figure 4.9 where the unfiltered cross sections are compared with the filtered ones. The filtered profiles are narrower than the original data since the filtering is not generally valid close to the edges unless the data are padded appropriately.

For the stitching of real data in the results section, a simpler averaging filter was used in the form of a normalized 10×10 kernel with each element equal to unity. The normalization factor was therefore ×1/100. This corresponds to a standard filter available in commercial interferometry software packages [106]. The Gaussian filter used in this section was not used on real data as it was later discovered that there is a great deal of ambiguity in the ISO specifications for this type of filter [107]. Choosing the uniformly weighted averaging filter was deemed to be more compatible with the simpler filters in standard interferometry software, where Gaussian filtering

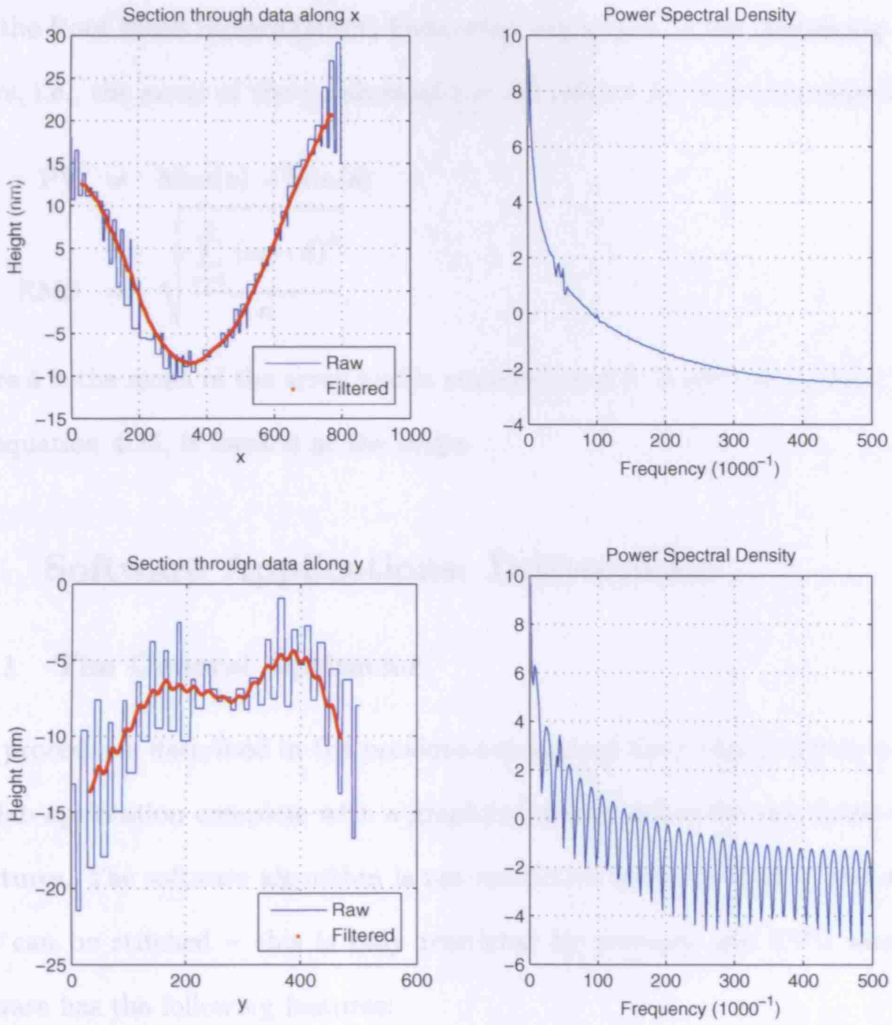


Figure 4.9: Comparison of the data cross-sections for the unfiltered and filtered stitched data. The filtered profiles clearly have the correct form, but without the noise generated by stitching. The filtered profiles are narrower due to the cut-off at the edges of the data patch.

is left for future work.

4.8.1 Statistics

The statistics computed for stitched surfaces are the Peak-to-Valley (PV) form error and the Root Mean Square (RMS) form error. Let **s** be a vector containing the form errors, i.e., the array of the n values of z in the plot in figure 4.6 on page 188.

$$\text{PV} = \text{Max}(\mathbf{s}) - \text{Min}(\mathbf{s}) \tag{4.26}$$

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^{n} (s_i - \bar{\mathbf{s}})^2}{n}}, \tag{4.27}$$

where $\bar{\mathbf{s}}$ is the mean of the array and is nominally equal to zero as the data centroid, see equation 4.25, is located at the origin.

4.9 Software Applications: Deliverables

4.9.1 The General Optimizer

The procedures described in the previous subsections have been programmed into a Matlab application complete with a graphical user interface for any shape of n sub-apertures. The software algorithm is not restrictive in the number of sub-apertures that can be stitched – this is only restricted by memory and CPU issues. The software has the following features:

1. A window to input the optimizer seed parameters.

2. A window to view the optimized transformation parameters.

3. A visual indication of the layout of the fiducial matrix (this is known as a spy plot in Matlab).

4. A graph showing the object function convergence to zero (the ideal value).

5. A graph showing the directional derivative values as a function of the iteration number.

6. 3-D plots of the fiducial points before and after optimization.

7. An optimization status display.

8. An extensible algorithm that determines how many sub-apertures to stitch from the fiducial matrix and initial transform parameters.

In addition to the above, the program opens separate windows displaying the stitched and correctly scaled measurement result which can be programmed to be filtered if necessary.

Figure 4.10 shows the typical layout of the GUI for the stitching software and the process parameters. The plot to the upper left displays the object function convergence to zero after each iteration and the upper center plot shows the magnitude of the directional derivative of the function at each iteration. The two lower plots display the positions of the fiducials before and after optimization (the positions of the fiducials are projected onto the x-y plane for these plots). The plot in the upper right is a spy plot of the fiducial matrix for this software. A spy plot is simply a plot of a matrix and has points where there are non-zero entries in the matrix. The box in the lower right of the window displays the optimizer status, displaying the number of iterations, the object function value and the termination status

The features of the window are common no matter how many sub-apertures are stitched. The buttons in the lower portion of the window are used to control the application. If simulated data is used each sub-aperture must be perturbed from its correct position. Selecting the 'Seed/Perturbation' button opens the seed parameters screen where the 6 initial parameters can be set for each sub-aperture. Similarly, when the data have been stitched, pressing the 'Optimized parameters' button brings us a similar window showing the stitching parameters for each sub-

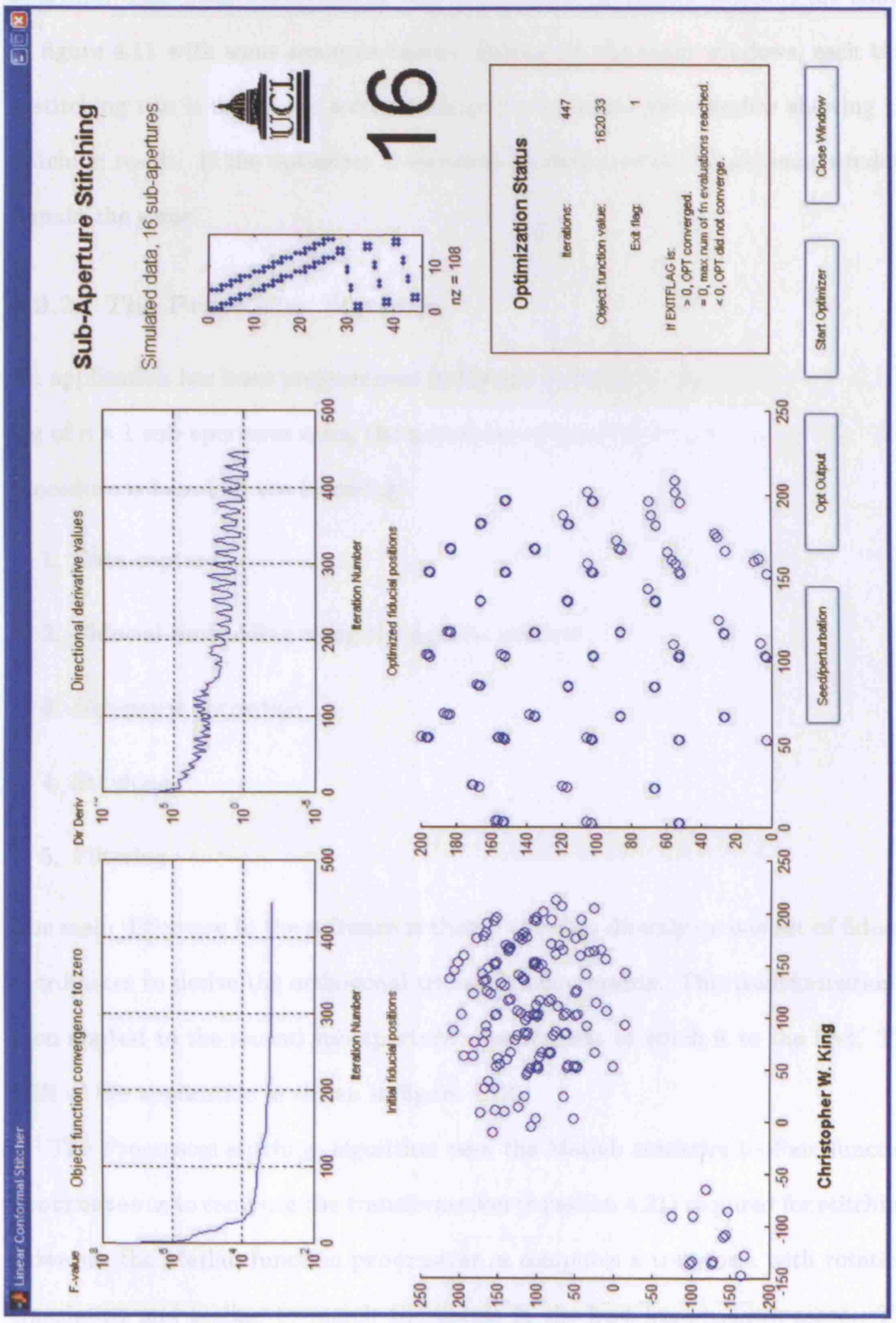


Figure 4.10: The stitching software main GUI showing the progress of a 16 sub-aperture stitch. See main text for an explanation of the plots.

aperture. The 'Seed/Perturbation' and 'Optimized' parameter windows are shown in figure 4.11 with some example values. Further to the main windows, each time a stitching run is made, the software outputs a separate plot window showing the stitching result. If the optimizer is switched to `fminsearch` the software windows remain the same.

4.9.2 The Procrustes Software

An application has been programmed in Matlab to facilitate the Procrustean stitching of $n \times 1$ sub-apertures using the principles outlined in the previous section. The procedure is based on the following:

1. Data capture
2. Fiducial centroiding using the annular method
3. Geometric correction
4. Stitching
5. Filtering

The main difference in the software is that it operates directly on one set of fiducial coordinates to derive the orthogonal transformation matrix. This transformation is then applied to the second sub-aperture measurement to stitch it to the first. The GUI of the application is shown in figure 4.12.

The Procrustes stitching algorithm uses the Matlab statistics toolbox function `procrustes.m` to compute the transformation (equation 4.21) required for stitching. However, the Matlab function `procrustes.m` computes a transform with rotation, translation and scaling to match the points in the best least-squares sense. The scaling is unacceptable in this application as has been stated in section 3.1.1. The algorithm is therefore used to give the best transformation without scaling being applied. The algorithm returns three objects;

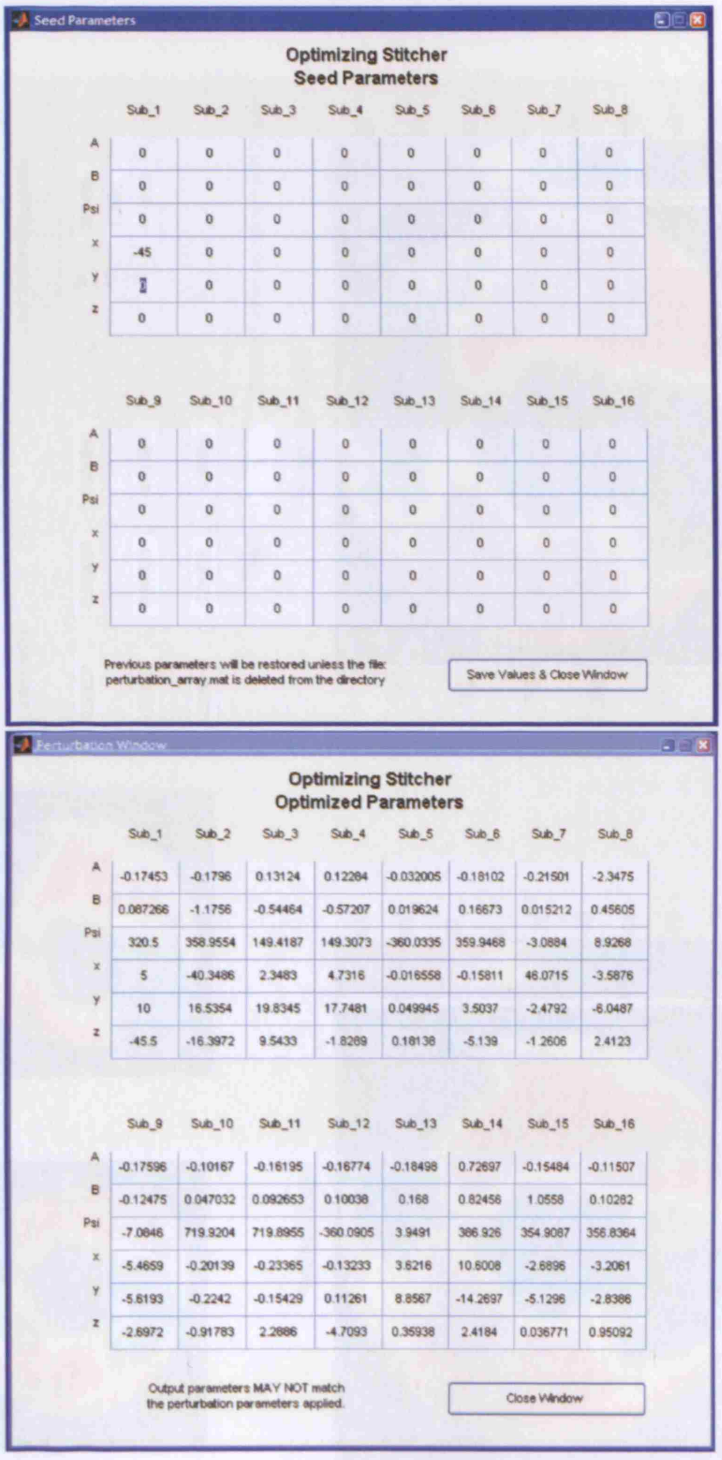


Figure 4.11: The stitching software parameter windows. The seed window (upper) is used to input the optimizer seed or the transformations to be applied to the simulated data before stitching back together, i.e., the simulated data have to be unstitched first. The lower window is the actual output from the optimizer and is used for both simulated and real data.

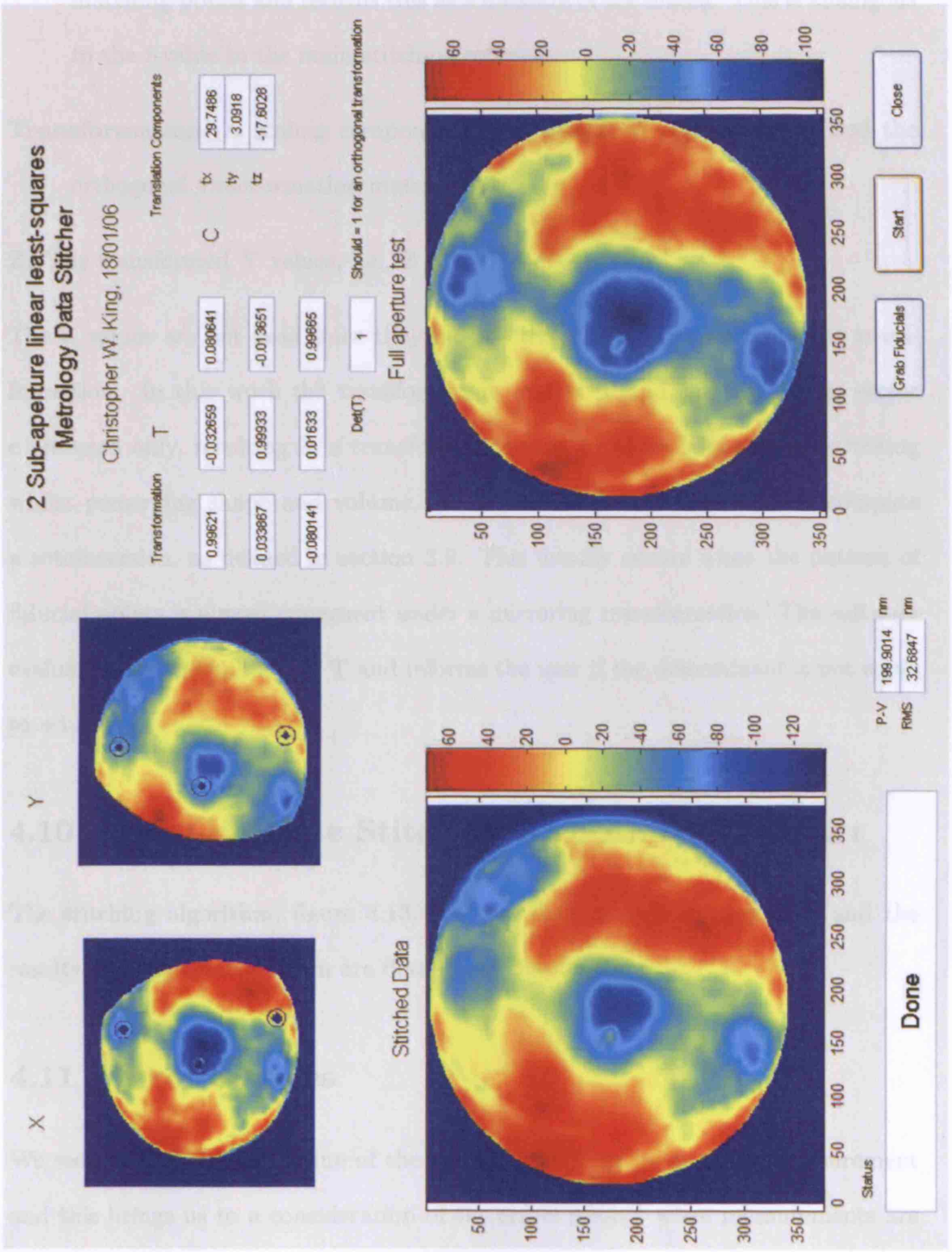


Figure 4.12: The Procrustes stitching software main window.

Measure of fit: The algorithm computes the sum of the squared errors between matching points and returns this as a measure of the match. This is analogous to the f-value in the main stitching software.

Transformation: A scaling component, b, a translation component, $\mathbf{c}$, and the orthogonal transformation matrix, $\mathbf{T}$.

Z: The transformed $\mathbf{Y}$ values, i.e., $\mathbf{Z} = b\mathbf{YT} + \mathbf{c}$.

The $\mathbf{Z}$ values are not used since they include the all the parameters for the transformation. In this work the transformation matrix $\mathbf{T}$ and the translation vector $\mathbf{c}$ are used only, resulting in a transformation that provides least-squares stitching whilst preserving shape and volume. It is possible for the algorithm to compute a rotoinversion, as defined in section 3.2. This usually occurs when the pattern of fiducial points is almost congruent under a mirroring transformation. The software evaluates the determinant of $\mathbf{T}$ and informs the user if the determinant is not equal to $+1$.

4.10 The Complete Stitching Algorithm: Flowchart

The stitching algorithm, figure 4.13, has now been completely described and the results of using the algorithm are detailed in chapter 6.

4.11 Error Sources

We seek to find the true value of the surface form error by stitching measurement and this brings us to a consideration of the errors present when measurements are made. There are many sources of error when making interferometric measurements for stitching, but there are a few which are more significant than the others.

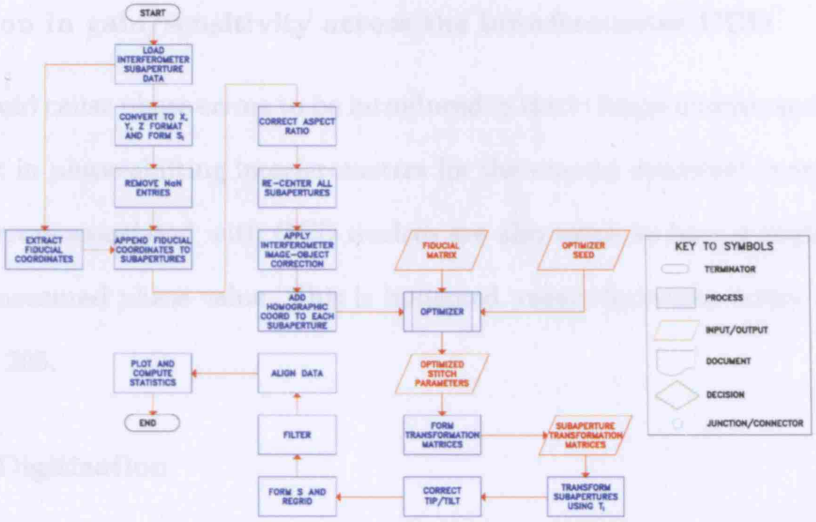


Figure 4.13: The complete flowchart for the optimizing stitching algorithm.

4.11.1 Random Errors

Environmental Errors

These include vibration, air turbulence, thermal gradients and mechanical drift. In this work, these errors have been estimated by taking repeated measurements of a part where the optical path in free air was less than 1 m. The RMS deviation of the measurement points from the mean surface was used as the estimate for these errors. This was estimated as approximately 4 nm. Allowing a little extra we can therefore reasonably expect errors in the range of 4 – 12 nm from this source.

Frequency stability of the source laser

These errors would manifest themselves as a scaling error in the phase values across the whole phasemap. Modern frequency stabilised lasers are extremely stable and the effects of this are taken to be negligible.

4.11.2 Systematic/Calibration Errors

Variation in gain/sensitivity across the interferometer CCD

This would cause phase errors to be introduced in static fringe interferometers. It has no effect in phase-shifting interferometers for the reasons described in section 1.6.2. Other errors associated with CCD devices are also taken to have a negligible effect on the measured phase value. This is budgeted under electronics errors in table 4.2 on page 205.

Phase Digitization

Modern interferometers digitize the interferograms using a CCD device and analogue to digital conversion. This places a limit on phase digitization. The Wyko 6000 system has a quoted resolution of $\lambda/1024$, corresponding to 0.6 nm at 633 nm. This is budgeted under electronics errors in table 4.2.

Phase stepping/modulation

Errors will occur if the phase is not stepped by exactly known amounts. Selberg states that modern phase-shifting interferometers perform very accurate calibration of the phase-shifting and that the residual errors are negligible. [108]. This is budgeted under electronics errors in table 4.2.

Fiducial centroiding errors

The errors in x- and y-centroiding come from imaging the surface with the interferometer. The imaging resolution of the optical system of the interferometer is given by the Rayleigh formula to a first approximation – see footnote to page 40. With D equal to the aperture of the interferometer at 150 mm (and equal to the diameter of the main collimating lens) we obtain an angular resolution of 5 μrad. In practice, this is further limited by the pixel resolution of the camera. On plano parts the

spatial resolution is limited to approximately 0.5 mm at full aperture. For spherical parts this figure depends on the radius of curvature, with shorter radii and lower $f/\#$ parts being better sampled.

The centroiding method described in chapter 2 uses the interferometer's phase recovery system to identify fiducial pixels. If insufficient phase modulation occurs on a given pixel between the 4 interferograms then the phase at that pixel is not assigned, and that pixel then belongs to a fiducial. Therefore any directional effects in the fiducial can cause a given fiducials pixel count to change between two different sub-apertures. Experiments in this work have found the error to be on the scale of 2 pixels on the CCD which can be mapped to the part scale by knowing the imaging magnification.

A further source of x-y error comes from geometric distortion as discussed under calibration errors.

The z-values are subject to smoothing out since the average values of the phase of an annulus of pixels is computed about the x-y centroid, furthermore, the z-error is slope dependent since it relies on the x-y centroid.

Reference surface errors

These errors result from the imperfect form of the reference surface. They can be calibrated out by having a master reference test surface with the same radius as the parts to be tested. Standard interferometer reference surfaces are normally specified to be flat or spherical to $\lambda/20$ or about 30 nm P-V at 633 nm.

Aberrated wavefronts change shape as they propagate so that the test and reference wavefronts are not the same shape when they interfere - even if the test surface is perfect. The effects are minimized in equal path interferometers. In unequal path interferometers the magnitude of the effect depends on the ratio of the radii of the test and reference surfaces - hence the requirement for a good master surface to

calibrate the errors. The RMS errors of such surfaces are generally $\lambda/50$ or better - about 12 nm. The form errors dominating the P-V are very often located at the outer edges of the reference surface and it is possible to mask out these areas in software if they are not required for use, thus improving the measurement.

Re-trace errors

Re-trace errors are a type of ray-mapping error. Slope errors on the test surface relative to the reference cause the rays not to retro-reflect. The resulting optical path difference is the aberration of the system. When testing spheres the test and reference surface radii should be closely matched since the magnitude of the re-trace errors increases in proportion to the ratio of their radii. For plano parts the test surface should be located as close as possible to the reference surface to avoid long return paths.

We can provide a simple estimate of the re-trace errors by considering a slope error on an otherwise perfect plane surface as shown in figure 4.14. We let the interferometer cavity length be equal to l and the slope error to be equal to θ. The length of the return path of the reflected ray to the reference surface is given by,

$$\text{PATH} = \frac{l}{\cos\theta}. \tag{4.28}$$

Therefore the optical path difference (OPD) caused by this slope error is,

$$\text{OPD} = \frac{l}{\cos\theta} - l. \tag{4.29}$$

For a typical optical surface which is flat to $\lambda/20$ P-V, we might encounter a change in height of 30 nm over 10 mm at an edge where most of the P-V usually occurs. This corresponds to a linear slope of 3 μrad. Inserting this into the equation 4.29 with l=0.5 m gives[1] an OPD of 0.45 nm. The distance along the reference surface between the point where the ray was launched and where the reflected ray intersects

[1] l = 0.5 is chosen to be slightly longer than the paths used in this work.

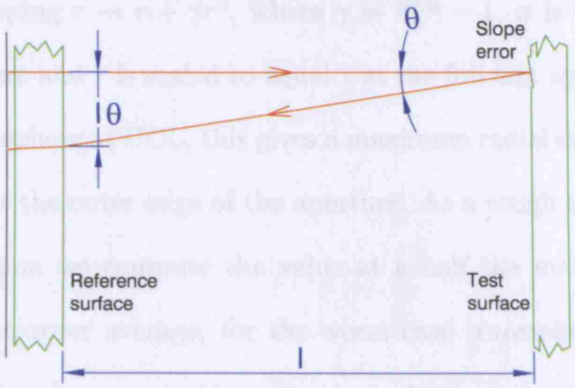


Figure 4.14: Estimating the re-trace errors by considering a portion of a plane surface with a slope error.

the surface is given by $l\theta$ and in this case equates to 1.5 μm. After the imaging system demagnification is accounted for, this dimension is much smaller than the size of the interferometers CCD pixel size, and it is therefore seen that the ray does not interfere with a different pixel.

These arguments apply very well when the fringe patterns are nulled, and Selberg notes that re-trace errors should be considered when the interferogram contains high fringe densities[loc cit.] - precisely when the slope errors are much larger, and as required in static-fringe interferometry. Long test cavity lengths also worsen the re-trace errors in unequal path interferometers, but the problems here will generally be very small compared to errors caused by vibration and turbulence. A similar calculation and numbers apply to spherical testing.

Imaging Distortion

The imaging system of the interferometer will introduce distortion errors which generally stretch the phasemap unevenly [109], especially when spherical testing is performed. Full calibration can be performed with a reference object containing calibration marks and is a complicated and lengthy process.

A very approximate formula for the radial distortion by the reference element is

given by the mapping $r \rightarrow r + \gamma r^3$, where $\gamma = \frac{\sin\alpha}{\alpha} - 1$, α is the half angle of the transmission sphere and r is scaled to equal 1 at the full test aperture [108]. For the $f/1$ transmission sphere at UCL, this gives a maximum radial distortion of 4% of the aperture radius at the outer edge of the aperture. As a rough estimate for the error caused by distortion we compute the value at a half the maximum radius of the aperture as a first-order average, for the worst case transmission sphere available ($f/1$). The fractional error at this radius is 1×10^{-2} of the radius value. Selberg notes that this distortion is also proportional to the phase slope, with the highest accuracy achieved for a null test. The distortion does not affect the phase values at null, as it is the interference pattern that is distorted. The distortion is therefore a geometric effect on the phase data as a whole, and is important when the phasemap needs to be mapped onto the part exactly as is required in stitching.

4.11.3 Budget

The errors previously discussed are collected and shown in table 4.2 below, as appropriate for the testing performed in this work. We can now begin to estimate the error on any one phase measurement, and as a 'worst' case, choose the situation of using an $f/1$ transmission sphere measuring a part used in chapter 6. The part has a radius of 284 mm and an aperture diameter of 200 mm.

It is possible for the phase errors to sum with the same sign, this would give:

$$E + F + I + R + M = 26.6 \text{ nm}. \tag{4.30}$$

This gives an estimated maximum error for the phase at any point on the phasemap, and it is unlikely that this will occur across adjacent pixels. The errors will be uncorrelated and as they are RMS errors we can root sum square to yield a better estimate of the phase errors:

$$\sqrt{12.0^2 + 0^2 + 1.3^2 + 12.7^2 + 0.6^2} = 17.5 \text{ nm}. \tag{4.31}$$

This corresponds to approximately $\lambda/37$ at 633 nm.

Error	Symbol	Effect	Magnitude / nm	
Environmental	E	$\lambda/50$	12.0	RMS
Laser frequency	F	$< \lambda/1 \times 10^6$	0	
Fiducial centroid: x-y		2 - 8 pixels	Distortion dependent	
Fiducial centroid: z			Slope dependent	
Electronics/instrument	I	$< \lambda/500$	1.3	RMS
Reference surface	R	$\lambda/50$	12.7	RMS
Retrace errors	M	$< \lambda/1000$	<0.6	RMS
Distortion			$\frac{\Delta r}{r} = 1 \times 10^{-2}$	

Table 4.2: Contributing errors to the stitched measurement result. Fiducial errors are those found in chapter 2.

The distortion estimation can be applied to arrive at the error in a fiducial's x-y coordinates. We consider a fiducial location somewhere at r = 0.5× the full aperture radius where the fractional error has been approximated by 1×10^{-2}. On the example part, a CCD pixel corresponds to approximately 1.3 mm, and the fiducial might be located at a radial distance of 50 mm (the part's full radius is 100 mm). The error in the real radial coordinate is $1 \times 10^{-2} \times 50 = 0.5$ mm. At 1.3 mm/pixel our error in the coordinate is approximately 0.65 pixels. We require the fiducial coordinates to be at least as accurate as this for the best possible stitching. We therefore choose 0.5 pixels as the target for the x-y centroiding. As shown in chapter 2, the current fiducial centroiding error is approximately 2 pixels and a factor of 4 improvement is therefore required. It is envisaged that this figure can be reached and even improved with a more modern interferometer and further work on distortion correction involving a more complete mapping of the imaging systems.

We estimate the fiducial z-error as follows. For a good part with a form error of $\lambda/20$ P-V, we could reasonably expect a slope error of 3 μrad, which, combined

with the 0.5 mm error in fiducial location, corresponds to 1.5 nm in z. The z-error is also a function off the phase error and we root sum square to obtain 17.6 nm as the target for centroiding in z. In the trials in chapter 2, the error was found to be much less than this.

This conservative analysis has been based on the experiments performed in chapter 6 for the test configurations used there. The next chapter digresses a little to describe the NPL/UCL swing-arm profilometer and the author's own contributions to this project.

CHAPTER 5

The NPL/UCL Swing-Arm Profilometer

This chapter comprises an overview of the NPL/UCL swing-arm profilometer and how it fits in with stitching metrology data. The chapter is presented as a related digression to the main stitching theme throughout the rest of the thesis on the basis that the swing-arm profilometer provides a suitable platform for a future fully automated stitching metrology instrument. Further to this, the author's own contributions, both in hardware and software, to the swing-arm profilometer are sufficient in themselves to be contained within their own chapter.

The concepts developed in chapters 2 to 4 show a method for stitching metrology data from overlapping sub-aperture measurements. The positioning between the interferometer and the surface under test can be arranged manually, by means of physically moving the part or the interferometer by hand into each of the test locations, or, the relative position can be controlled by a computer driven machine platform. The latter method has the distinct advantage in that the whole operation can be automated – pre-set interferometer–part relationships can be programmed

into the system based on the configuration of overlapping sub-apertures. The positional accuracy, unlike that of the QED system by Fleig et al. [loc cit.], need only be of the order of one millimeter since the physical fiducials provide the exact stitching points between sub-apertures. Furthermore, the optimizer could be seeded by the encoder outputs on the motion axes of the swing-arm profilometer.

The NPL/UCL swing-arm profilometer project started in 2002 after a review of how future ELT segments might be made and measured [30]. The profilometer forms stream (A) of the OMAM work described on page 69 [59]. The mounting of a wavefront sensor on the swing-arm profilometer would fully constitute a stitching metrology platform. Further work in strand (C) of the OMAM project described on page 69 deals with the techniques for attaching a wavefront sensor to the swing-arm profilometer [79].

5.1 Swing-Arm Profilometry

The idea of a swing-arm profilometer for optical testing is not new – it was first suggested by Anderson, Parks and Shao [110] and indeed used by Andersen and Burge in 1995 [111] at the Steward Observatory Mirror Lab to make measurements of ground and polished parts. A swing-arm profilometer consists of a length of rigid arm attached to a very precise rotary bearing such as an air-bearing. As the arm rotates on the bearing, the point at the end of the arm will trace out a circular arc in space, with a radius equal to the perpendicular distance from the axis of rotation to the end of the arm, denoted by l. Now, take a spherical convex part with radius of curvature equal to R. If the centre of curvature of the part is made to intersect with the arm's axis of rotation, and the arm tip is initially placed at the optical centre of the part, then the arm's end will trace out an arc that is exactly aligned on the surface of the part. By placing a contacting probe on the end of the arm, the part's deviation from spherical can be measured directly. The concept of

a swing-arm profilometer is shown in figure 5.1.

Any surface with a radius of curvature, R, larger than l, can be measured by tilting the rotation axis of the swing-arm according to,

$$\sin\theta = \frac{l}{R}, \tag{5.1}$$

where R is the radius of curvature and θ is the tilt angle of the arm, and ensuring that the arm's rotation axis intersects with the centre of curvature of the part under test. When $R = \infty$, $\sin\theta = 0$ and therefore $\theta = 0$, we are measuring a flat and the two axes never intersect. In all instances when equation 5.1 is met, the arm tip traces out a circular arc that sits on top of the surface under test.

A simple way to visualize this is that any ring of radius, l, where $0 < l \leq R$ can be placed upon any sphere or radius R – the ring represents the path traced out by the end of the swing-arm.

If the condition $\sin\theta = \frac{l}{R}$ is held then the arm path will be exactly coincident with the surface of the convex sphere, and the idea can be extended to any spherical surface be it convex or concave provided that the radius of the surface is greater than or equal to the arm length l.

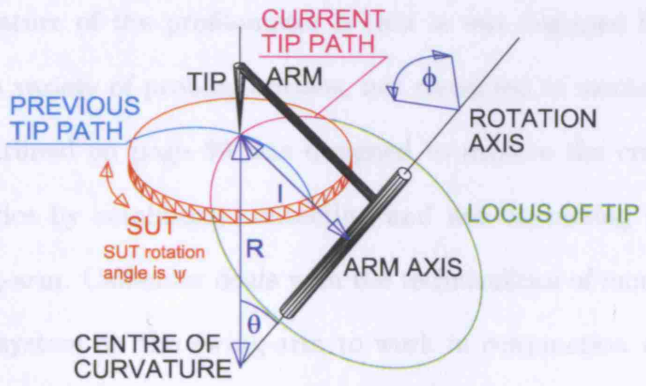


Figure 5.1: Swing-Arm Profilometer scheme. The green circle is the locus of all tip points and has a radius equal to l. Two possible tip paths are shown on the surface under test (SUT) which can be rotated about its axis as shown in red.

In practice mechanical constraints will limit the range of coverage in terms of R.

There are many possible mechanical configurations of a swing-arm profilometer and the only constraint is that the swing-arm condition, equation 5.1 must be satisfied.

A contacting metrology probe placed at the tip of the arm will give a direct measure of the aspheric departure with respect to the reference radius, R. This contacting probe, once set normal to the surface under test will remain normal as the arm is swung about its axis. A further advantage of a swing-arm profilometer is that one only has to change the mechanical (or optical) probe to increase or decrease the maximum aspheric departure that can be accommodated.

5.2 The NPL/UCL Design

The National Physical Laboratory joined the OMAM project in 2002 and were funded by the DTI under the national length programme to produce a prototype swing-arm profilometer that could be used to measure optical surfaces up to 1 m in diameter. This coincided with the requirements of the UCL basic technology project to produce 1 m diameter demonstration ELT mirror segments in ULE glass and in silicon carbide.

A novel feature of the profilometer is that is was designed from the outset to be used with a variety of probing options, not restricted to mechanical probes only. Strand (C) outlined on page 69 was designed to explore the creation of a hybrid metrology device by combining contacting and non-contacting metrology devices onto the swing-arm. Callender deals with the technicalities of mounting a wavefront measurement system on the swing-arm to work in conjunction with a mechanical probe [79].

One of the conclusions from this work was that a direct link must be made between the wavefront sensor data and the contacting metrology probe for the data to be useful. This could be significant for aspheric work, where the radius of curvature change across the part must be tracked before stitching can be performed. The Zygo

verifire asphere [loc. cit.], although purely an optical metrology platform, achieves this for axially symmetric parts by measuring annular zones on the part. These are measured by an Fizeau interferometer and transmission sphere at the points where the test wavefront tangents match the surface under test. As the part is moved along the optical axis relative to the interferometer, the tangents will match at different annular zones on the surface under test. At each measurement zone a precise measurement of the part's distance from the interferometer is taken using a displacement measurement interferometer. The displacement data is used to correctly map the positions of the annular zones after all measurements have been made. In this way, the Verifire asphere can track the curvatures at different zones and reconstruct the surface from of the part under test.

The NPL/UCL swing-arm profilometer was built by modifying an existing CMM manufactured by Fishers Aitken-Gesellshaft, consisting of a granite base, x-y translation stage and a z-axis translation stage mounted vertically on the x-y stage. The original granite base and x-y stage were retained. The x-y stage provides 400 mm of travel in each direction and a nominal resolution of 100 nm and allows the main air-bearing's axis of rotation to be moved in the x-y plane. This is used to intersect the main air-bearing's rotation axis with the centre of curvature of the surface under test.

The original z-stage of the CMM has been replaced by the air-bearing mounting. The arm is slung from the air-bearing in a cradle and trunnion system where this has the advantage that the arm can be held horizontal when in the mid-position – see figures 5.3 and 5.4, for example. The arm is composed of a 1220 mm length of hollow rectangular section alumina tube chosen for its high stiffness and low thermal expansion coefficient. The servo-controlled main air-bearing is a Professional Instruments BLOCK-HEAD 10" diameter precision air-bearing with 25 nm axial and radial run-outs. The main spindle is connected to a Heidenhain rotary encoder

with a resolution of 0.36 arc sec.

Stub axles on the sides of the air-bearing allow the tilt to be finely controlled via a side mounted bell crank mechanism driven by a standard micrometer. This arrangement allows sub-arc second adjustment of θ. Finally, parts to be measured are mounted on a servo-controlled rotary air-bearing table with a two-axis centering stage.

The end of the arm contains an optical breadboard to which several probing arrangements can be attached and full details of this aspect can be found in Callender [79].

Efstathiou [59] contains full design and construction details of the swing-arm profilometer. More brief details can be found in Callender [112], King [113] and Lewis [114].

5.2.1 Arm Length: fixed or variable?

The arm length was the subject of much debate at the design stage and two alternative arrangements were proposed. One was to have a fixed arm length chosen to accommodate most of the optical components that were envisaged to be encountered; this is achieved by consideration of equation 5.1 and the diameter of the part. The other, suggested by the author, was to have a continuously variable arm length by using a fixed piece of alumina mounted in a kinematic clamp and hung under the main arm-rotation bearing. This would allow very short arm lengths, and hence small radii of curvature to be measured (<1 m). However, concerns were made about the stability of such systems on the nanometre level and the final solution adopted was to have an arm fixed in a cradle mounted asymmetrically about the long-axis of the arm. This way, the arm mounting is kept simple and allows for two arm lengths to be used by turning the arm around.

5.2.2 The Swing-Arm Coordinate System

The author of this work provided the coordinate system transformations for the swing-arm profilometer after considering how to find the coordinates of the centre of field of an optical metrology device attached to the swing-arm. The analysis is broken down by describing in Cartesian form the motions of the swing-arm probe tip or axis in terms of the swing-arm coordinates. The coordinate system of the swing-arm profilometer uses the following variables:

- θ: the tilt angle of the swing-arm profilometer.
- R: the base radius of curvature of the part under test.
- l: the effective arm length from the contact probe null position at the part vertex to the swing-arm rotation axis, measured perpendicular to the swing-arm rotation axis.
- ϕ: the angle of the swing-arm from the zero position. The zero position is defined when the probe tip contacts the part vertex.
- ψ: the angle of the rotary air-bearing table. This is therefore the part rotation angle from the zero position defined by the rotary table.
- $\mathbf{z}$: the contact probe measurement value in the form $(0, 0, z)^{\mathrm{T}}$.

The *swing-arm reference sphere* is also defined in terms or θ, R, l, ϕ and ψ, and corresponds to the space spanned by swinging the arm and rotating the rotary table for a given θ, R and l in a coordinate frame with the origin at the test part's centre of curvature.

Any point on the surface under test can be represented in swing-arm coordinates by θ, R, l, ϕ, ψ and $\mathbf{z}$. The conversion from swing-arm coordinates to Cartesian coordinates is accomplished by a matrix transformation derived as follows.

Let **T** be the set of all points lying on the swing-arm reference sphere for a given θ, R and l. Then **T** can be described in terms of the angles and distances that are the key parameters of the swing-arm.

The probe tip traces out a circle in space that is tilted about the centre of curvature of the part under test. This circle - the tip path - can be defined parametrically in terms of the arm sweep angle, ϕ, and the other swing-arm variables. Let the circle be initially centred at the point $(0, 0, R\cos\theta)$, have a radius equal to l and be initially parallel to the x-y plane. The part vertex is located at $(0, 0, R)$ The points on this circle are located in Cartesian space by:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} l\cos\phi \\ l\sin\phi \\ R\cos\theta \end{pmatrix}. \tag{5.2}$$

This circle is rotated by an angle θ about the y-axis such that the point where $\phi = 0$ is coincident with the part vertex. For a rotation about the y-axis the following transformation is used:

$$\mathbf{R}_y = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \tag{5.3}$$

to yield the set of points on the tip path coincident with the reference sphere:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} l\cos\theta\cos\phi + R\sin\theta\cos\theta \\ l\sin\phi \\ -l\sin\theta\cos\phi + R\cos^2\theta \end{pmatrix}. \tag{5.4}$$

In order to locate the points in space fully, the part rotation angle ψ must be considered. Rotation by the angle ψ corresponds to rotating the air-bearing table which proceeds about the z-axis. Thus for a rotation of angle ψ about the z-axis,

$$\mathbf{R}_z = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.5}$$

Equation 5.4 is multiplied by this to give an expression for the Cartesian coordinates of a point relative to the centre of curvature of the part under test in terms of l, R, θ, ϕ and ψ:

$$\mathbf{T} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} l\cos\theta\cos\phi\cos\psi - R\sin\theta\cos\theta\cos\psi - l\sin\phi\sin\psi \\ l\cos\theta\cos\phi\sin\psi - R\sin\theta\cos\theta\sin\psi + l\sin\phi\cos\psi \\ -l\sin\theta\cos\phi + R\cos^2\theta \end{pmatrix}. \quad (5.6)$$

$\mathbf{T}$ forms the basis for turning probe data in the swing-arm system into real surface data in the Cartesian system. The mechanical probe must be normal to the surface under test for the swing-arm to work well. The probe's normality is assumed for the most basic calculation of the actual surface form. If the probe is off-normal then some measurement must be made and the analysis altered.

When the probe's motion-axis is normal to the reference sphere then the probe tip travels along this normal at the point under examination according to the actual form of the surface under test. To compute the coordinates of the surface under test at a given swing-arm configuration, the probe reading, $\mathbf{z}$, is multiplied by the reference surface normal unit vector and this quantity is added to the Cartesian coordinate of the reference surface defined by equation 5.6. The origin of the coordinate system was conveniently placed at the centre of curvature of the surface under test. Surface normals can therefore be computed by finding a unit vector in the direction of $\mathbf{T}$ in equation 5.6:

$$\hat{n}_{(x,y,z)} = \frac{\mathbf{T}}{|\mathbf{T}|}. \quad (5.7)$$

Therefore for any measurement we can convert to Cartesian coordinates by evaluating

$$\mathbf{S} = \mathbf{T} \pm \left(z \cdot \frac{\mathbf{T}}{|\mathbf{T}|} \right), \quad (5.8)$$

using addition or subtraction depending upon the definition of the probe motion z. Here, $\mathbf{T}$ is a point on the swing-arm reference sphere. This must be added

to the vector describing the probe tip's position to obtain the coordinates of the surface under test. Since the probe is always normal to the surface, it moves along the direction given by equation 5.7. Therefore multiplying the probe's reading from zero by equation 5.7 gives the vector to be added to the reference sphere coordinates to give the coordinates of the surface under test.

To remove the z-component of the sphere and thus obtain a projection of the surface onto the x-y plane we compute:

$$\mathbf{S}_{\mathrm{REF}} = \mathbf{S} - \begin{pmatrix} 0 \\ 0 \\ -l\sin\theta\cos\phi + R\cos^2\theta \end{pmatrix}. \tag{5.9}$$

A Matlab program was written to check the validity of this procedure that could also evaluate the surface normals for any configuration of the swing-arm profilometer. Sections of the program were re-used to transform real data from the swing-arm profilometer. The graphical user interface of the software is shown in figure 5.2.

5.2.3 Calculation of the Sweep-Angle Required

ϕ has been defined in section 5.2.2 as the angle of the arm from the mid-position. The maximum arm-sweep angle must be known for each part tested such that a procedure can be established to automatically measure the part. Appendix E.1 shows a method used to calculate this angle for any given part. To compute the maximum sweep angle, we evaluate

$$\phi_{\mathrm{sweep}} = 2\sin^{-1}\left(\frac{C}{2l}\right), \tag{5.10}$$

where,

$$C = 2\cdot\sqrt{\left(\frac{D}{2}\right)^2 - \left(\frac{R\cos\left(\sin^{-1}\left(\frac{D}{2R}\right)\right) - R}{\tan\theta}\right)^2}, \tag{5.11}$$

and D is the part diameter and all other variables are as defined in section 5.2.2.

A LabView routine was written to compute the maximum sweep angle for use in the swing-arm profilometer control software.

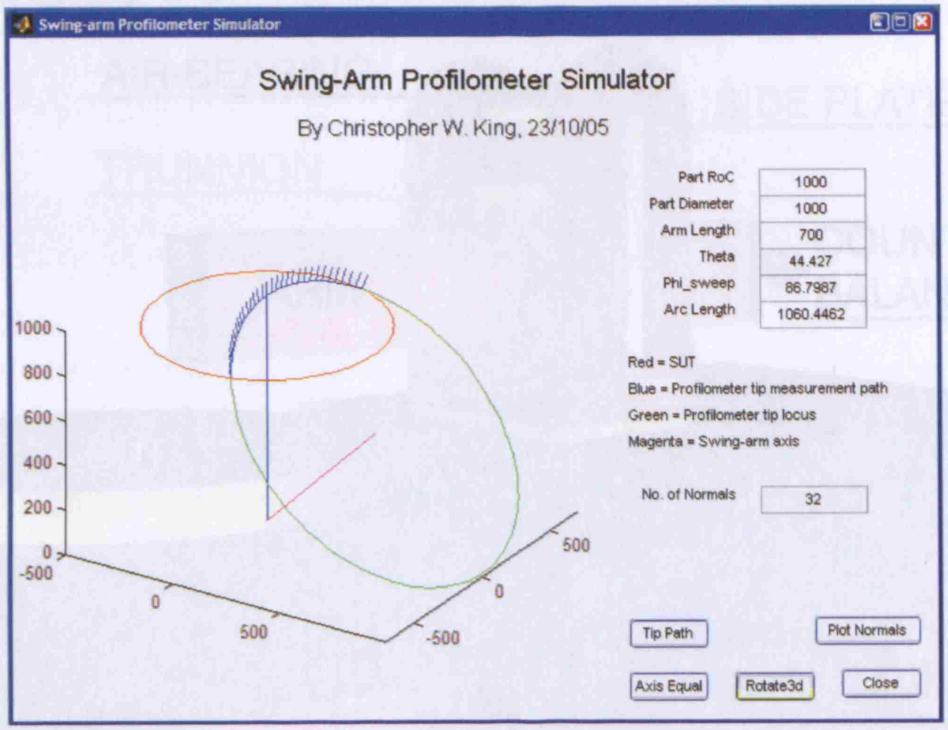


Figure 5.2: The SAP simulator software front-end showing the SUT, the probe tip locus, the swing-arm rotation axis, the centre of curvature and the surface normals.

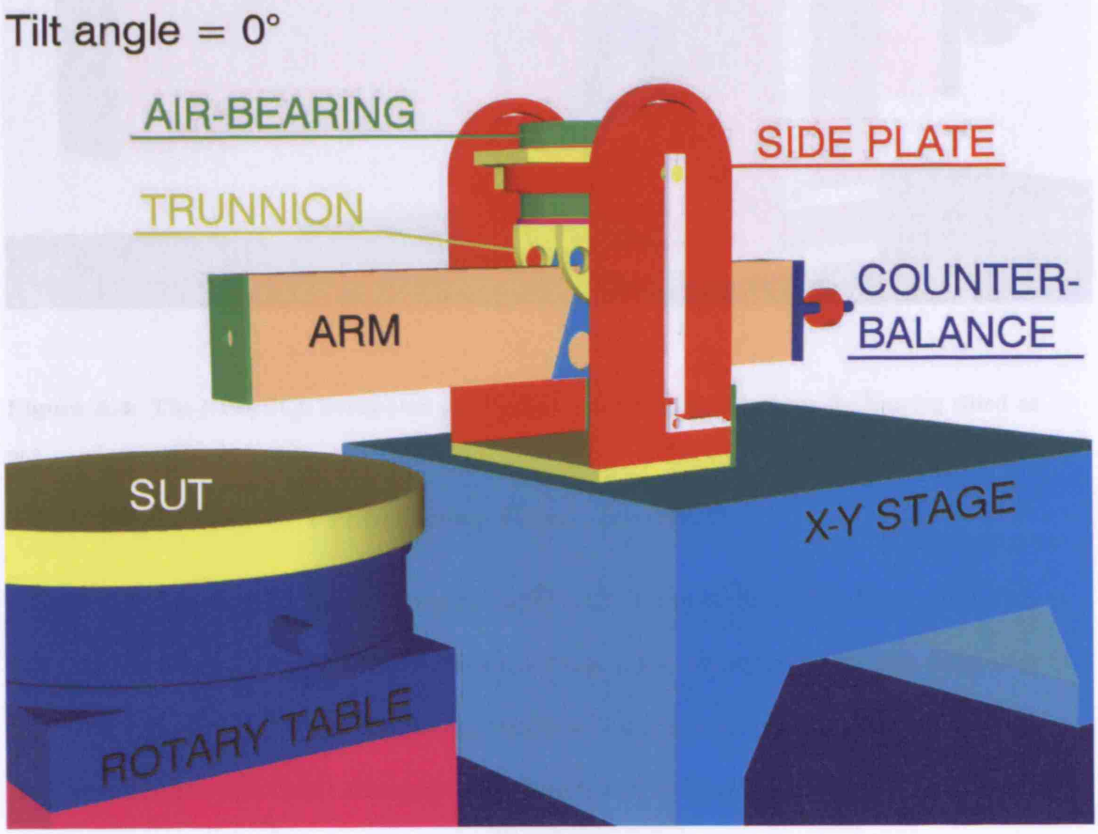


Figure 5.3: The NPL/UCL swing-arm profilometer main components with the bearing tilted at 0° .

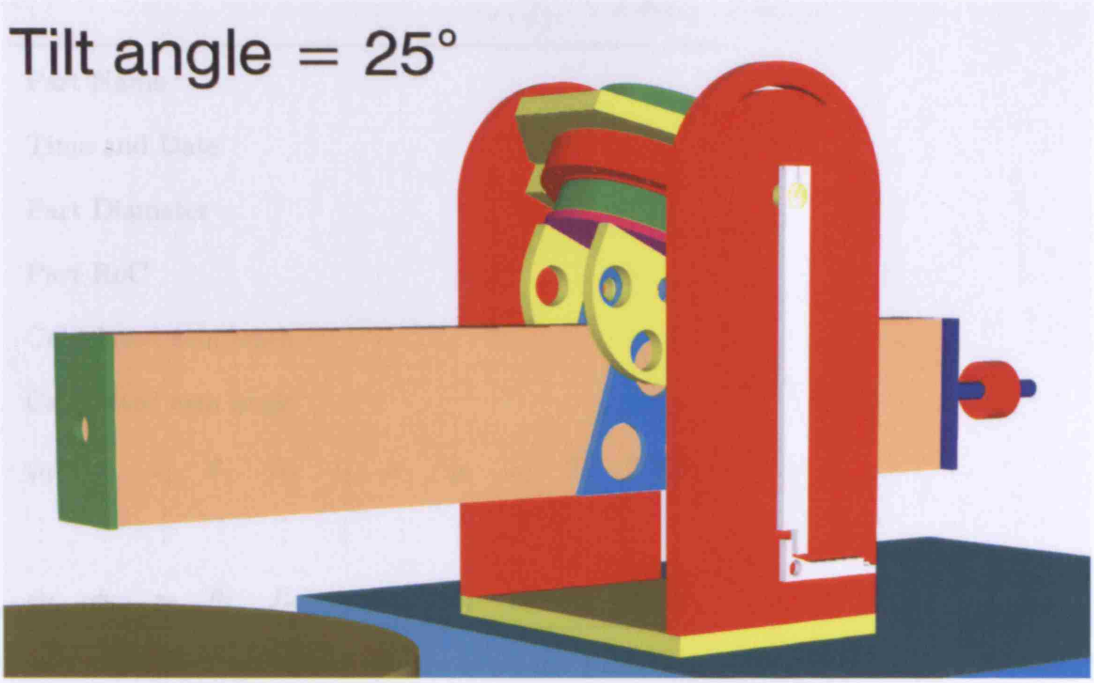


Figure 5.4: The NPL/UCL swing-arm profilometer main components with the bearing tilted at 25° .

5.2.4 Recovery of Probe Data From The SAP

All control software for the swing-arm profilometer is written in LabView, and during a measurement process the data shown in table 5.1 are recorded by the computer LabView software. The file format contains a header with user-defined part information and then the usual swing-arm variables and environment monitoring values. The arm length is currently calculated from the geometry of the swing-arm and the value of θ is calculated using a sine bar and gauge blocks. A precision sine bar is placed on top of the tilted swing-arm air-bearing and gauge blocks are placed under one end of the sine bar to make it horizontal, as measured by a very precise electronic spirit level. When the spirit level reads zero, the length of the sine bar and the thickness of the gauge blocks can be used to calculate θ as shown in figure 5.5. All other variables in the output data are measured from encoders or sensors by the acquisition hardware, and the file format can be read into Matlab for subsequent

Part Name										
Time and Date										
Part Diameter										
Part RoC										
Calculated arm lenth										
Calculated arm angle										
ψ_1	ϕ_1	z_1	θ_1	R_1	l_1	x_1	y_1	z_1	T_1	$\%RH_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
ψ_i	ϕ_i	z_i	θ_i	R_i	l_i	x_i	y_i	z_i	T_i	$\%RH_i$

Table 5.1: Output data format from the swing-arm profilometer with only a contacting metrology probe attached. The T_i record the temperature and the $\%RH_i$ record the relative humidity.

analysis and plotting as shown in figure 5.6.

The initial contacting probe used on the swing-arm profilometer was a Solartron LE12 probe with a range of 12 mm and a resolution of 12.5 nm. The probe is a linear encoder type utilizing an interferometric technique and a graduated scale to perform measurements.

5.2.5 First Data from the SAP

A 640 mm diameter concave sphere was mounted on the swing-arm profilometer for initial trials. The mirror has a radius of curvature of approximately 6096 mm. The mirror was placed on the rotary table and centred using the signal from the Solartron probe. Twelve arcuate scans were made across the mirror, each separated by 30° . Further to this, 10 concentric scans were made to give measurement paths on the mirror resembling a spider's web. The output data were in the form shown in table 5.2.4. The data in the SAP coordinates were transformed according to equation 5.8 and plotted in Matlab, to verify the transformations, however, the full

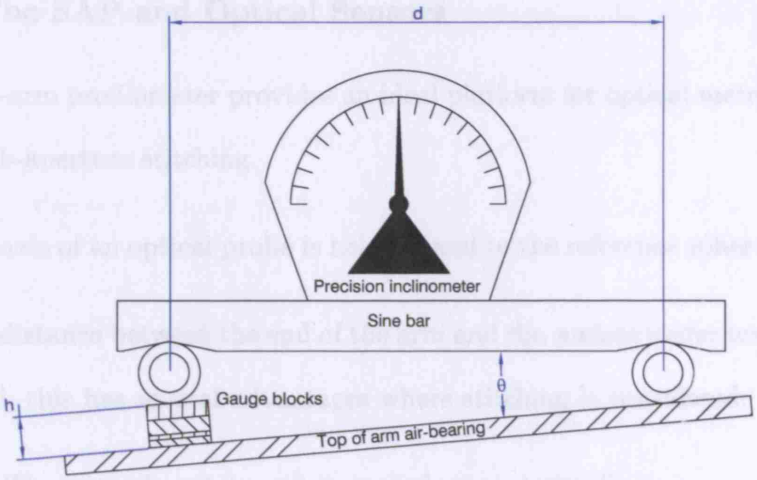


Figure 5.5: Using a sine bar and Johanssen gauge blocks to set the swing-arm bearing tilt angle. Here, the height of the gauge blocks to use is given by $h = d \sin\theta$.

analysis of the swing-arm data remains a project for future work as no resources were available for this task.

Plots of the data are shown in figure 5.6.

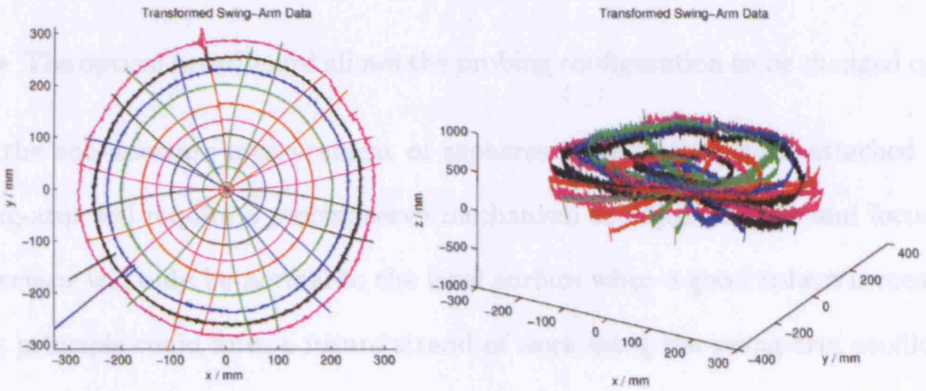


Figure 5.6: Transformed swing-arm profilometer data for the initial trials. Left: Plan view of the transformed data. Right: oblique view of transformed data highlighting the tip/tilt errors present in the instrument.

5.2.6 The SAP and Optical Sensors

The swing-arm profilometer provides an ideal platform for optical metrology devices and for sub-aperture stitching.

- The axis of an optical probe is held normal to the reference sphere at all times.
- The distance between the end of the arm and the surface under test is relatively small, this has several advantages where stitching is considered:
  - The optical path length in turbulent air is small.
  - In principle, a weak diverging transmission sphere could be used on an interferometer mounted on the arm to perform interferometric measurements at the virtual confocal point.
- The positions of the arm and rotary table are known from the read-out encoders and can thus be used as a seed for the stitching optimizer.
- The optical breadboard allows the probing configuration to be changed quickly.

For the sub-aperture measurement of aspheres a wavefront sensor attached to the swing-arm will require a precise servo mechanism to adjust tip/tilt and focus since the sensor will only be normal to the local surface when a good sphere is measured. This principle could form a future strand of work using the swing-arm profilometer and the stitching principles developed in this work.

The areal measurement range of the swing-arm profilometer is greatly increased when a wavefront sensor is used. Rather than measuring along an arc segment which is for all intents and purposes infinitely thin, a true areal sensor can give total surface coverage with a finite number of arm scans.

A hybrid system, combining contacting or non-contacting probe measurements with full 3-D surface measurements, allows for the possibility of traceable full-aperture stitching measurements. The probe system can be used to establish a

base radius in conjunction with accurate measurement of the swing-arm length and tilt angle. It is then possible to marry the probe data with the wavefront sensor data to provide a degree of traceability to the measurement result [79]. However, as discussed in the final chapter, much work remains to be completed to achieve this.

5.3 Further Work

Following the trials of the swing-arm profilometer and the assessments made of the performance [59] the following points were noted regarding the swing-arm:

1. The rotary air-bearing table is not sufficiently stiff and degrades the performance of the whole profilometer, causing wobble at the edge of the rotary table on the level of 1.0 μm but depending upon the position of the centre of mass of the load on the top surface. A new table or stiffer bearing would help to alleviate these problems. This could be further improved by the addition of interferometric monitoring of the table surface.

2. The data analysis software requires writing.

3. Software to automate the servoing of the wavefront sensor needs to be written.

4. To provide full traceability, the rotary table, the main arm rotation air-bearing and the tilt angle need interferometric monitoring to measure changes on the level of a few nm.

Items (1), (2) and (4) above are self explanatory. If we are to imagine the extension of the stitching software for use on the swing-arm profilometer in the future, we must consider the servoing of the interferometer to obtain a null measurement before we can even begin to consider the task of monitoring the radius of curvature changes of parts across the aperture.

Servoing in item (3) can be achieved automatically and efficiently by consideration of the interferogram and a simple idea is presented here.

Servoing the Interferometer

The best quality interferometer measurements are made as close as possible to null. This is when the test surface retroreflects the test beam back down the same path over which it arrived. A true null measurement is achieved for a perfect interferometer cavity when no fringes are visible in the interferogram - the phase of the interference pattern is constant because the optical path difference between the reference and test surfaces is constant. To achieve the best null test on a non-perfect cavity requires the number of fringes to be minimized. In a manually adjusted interferometric optical test the operator manually tilts and translates the test surface whilst observing the interference pattern until the number of fringes is minimized.

This process can be automated by consideration of the 2-D Fourier Transform (FT) of the fringe pattern. The 2-D FT will reveal the frequency and direction of the interference fringes, albeit with a 180 degree ambiguity. In the case of a perfect null test, the 2-D FT results in a sharp sinc function centred at the origin (assuming that the aperture is circular). For a tilted part, the 2-D FT is composed of two sinc functions symmetrically placed about the origin. The angle subtended between the peaks, axes and origin reveals the angle of the fringe pattern, and the radial distance of the peaks from the origin reveals the number of fringes (frequency).

A 2-D FT of the interference pattern can be computed very quickly and then analysed using the same procedures used in the automatic fiducial centroiding software of chapter 2. After centroiding the spots, a servo signal can be generated to be sent to the interferometer tip/tilt stage. The software could easily be made closed loop to servo the interferometer until the 2-D FT of the interference pattern is composed of a single peak at or near to the origin. The process is shown in figure 5.7. The top two plots show a non-null interference pattern with 5 waves of tip/tilt (10 fringes). The position of the spots on the right-hand plot is used to derive a servo signal that moves the interferometer. After two or three iterations, we might now

have the fringe pattern occupying the middle-left image of figure 5.7, where only tilt remains. Finally, in the bottom two images of the figure, the best null interferogram is shown along with the 2-D FT where the two peaks are almost located at the origin. A similar procedure of examining the 2-D FT of the fringe pattern can be used to remove defocus errors from the alignment.

This method mimics the way that many interferometer alignment modes work. Most modern interferometers switch an auxiliary lens into the beam to image the point source back onto the CCD camera. The point source image is reflected off the test surface and any tip/tilt will place an image of the spot onto the CCD device, but displaced from the optical axis which is usually marked on the auxiliary lens. The operator adjusts the tip/tilt controls until the spot is coincident with the optical axis. The extra lens effectively Fourier transforms the optical field and propagates the image of the point source onto the CCD device, which is equivalent to finding the 2-D FT of the fringe pattern.

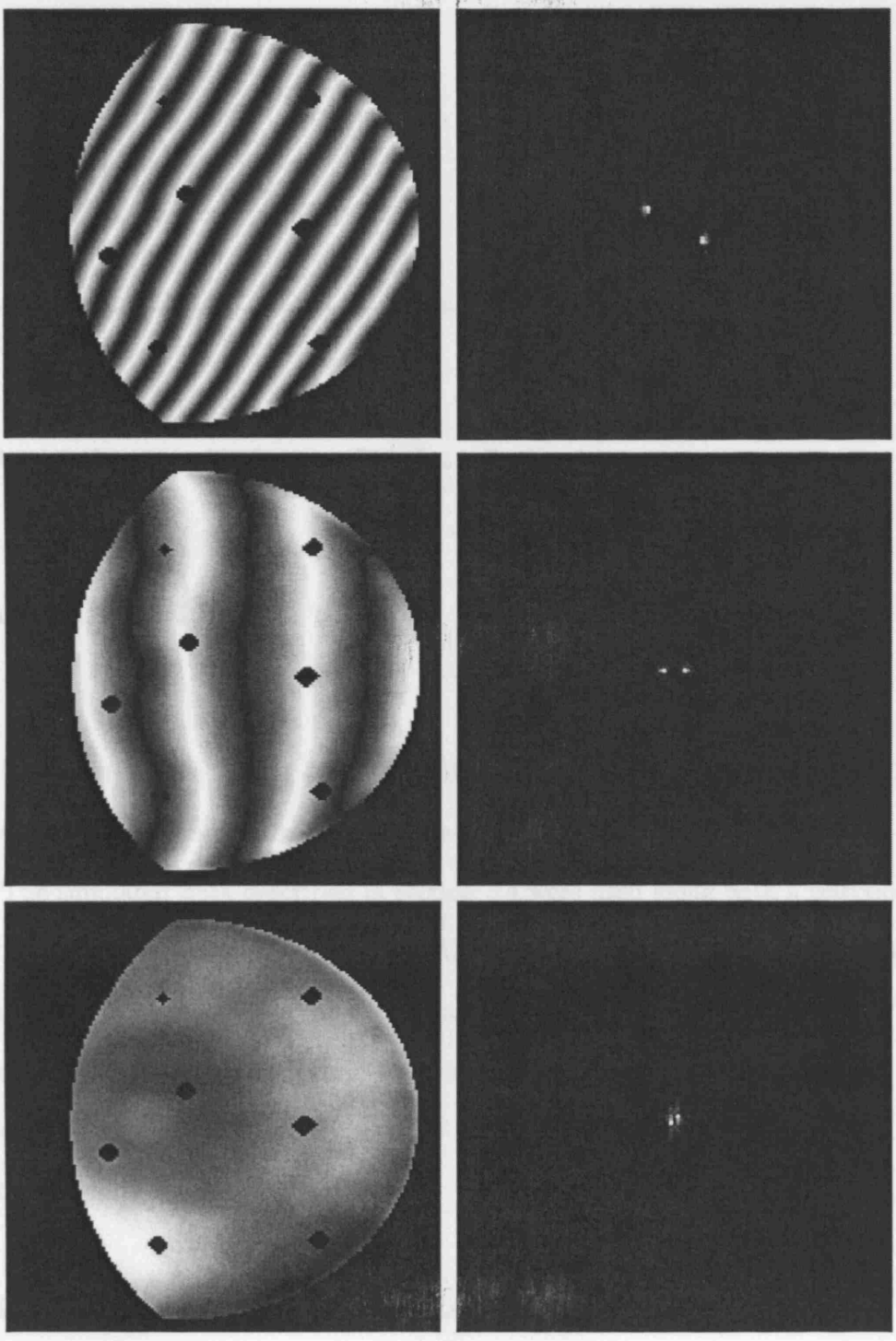


Figure 5.7: Left: simulated interferogram; Right: Fourier transform. Top: Real data with 5 waves of tip and 3 waves of tilt. Middle: 2 waves of tip and 0 waves of tilt. Bottom: Fringes nulled

CHAPTER 6

Testing The Stitching Algorithms

We now turn to the testing of the stitching algorithm developed in chapters 2–4. Tests were carried out using simulated and real surfaces.

The simulated data described in section 3.4 were used along with a selection of real flat and spherical parts available at UCL. Aspheric parts were not tested since this would have required an inherently more complex extension to the theory and a very precise radius of curvature measurement for which neither the budget nor timescales were available. A proposal for testing aspheric parts is contained in the section on further work.

6.1 1×2 Sub-Aperture Stitching: Simulated Data

6.1.1 The Procedure

The first test involved using two sub-apertures from the simulated data set. In theory, with no rounding errors and exact trigonometric functions is should be possible

to stitch the sub-apertures with no errors. However, with finite precision arithmetic, including roundoff and truncation errors [115], stitching will never be exact with simulated data.

An important feature of the software here is that only one transformation is applied to the sub-apertures, not a cumulative sequence of several, one for each iteration. If one transformation were applied at each iteration then the rounding and errors due to the finite float precision would be cumulative and after several hundred iterations could become very significant. These errors are thus avoided by applying only one optimal transformation matrix to each point in a sub-aperture.

Sub-apertures 1 and 2 were stitched with 3 fiducials. Here, there are no random errors – the data sets are fixed and definite and the transformations are applied to a precision limited only by the Matlab maths routines, which is determined by EPS, the distance between 1 and the next largest floating point number, and the trigonometric routines.

The following outlines the steps taken:

1. The simulated sub-apertures were perturbed from their stitched (normal) positions by applying transformations, $\mathbf{T}_i$, with random parameters, $\mathbf{t}_i$.

2. Three fiducial points were chosen in the data according to the scheme shown in section 3.4.

3. The fiducial matrix was composed and the data fed into the optimizer for stitching.

4. The optimizer was run until the target parameters were met.

5. The data were regridded and output for comparison with the original sub-apertures after tip/tilt correction.

The fiducial matrix for the simulated data is constructed by noting the columns in $\mathbf{S}_i$ where the fiducials are located according to the technique described section 3.3.4.

For example, the coordinates of fiducial 1 in sub-aperture 1 are located in column 3764 of $\mathbf{S}_1$. Similarly, fiducial 2's coordinates are located in column 4718 in $\mathbf{S}_1$, and in column 1218 in $\mathbf{S}_2$. Using these principles, the fiducial matrix for this stitching is simply given by:

$$\mathbf{F} = \begin{bmatrix} 3764 & 264 \\ 4718 & 1218 \\ 3714 & 214 \end{bmatrix}. \tag{6.1}$$

$\mathbf{F}$ indicates that each of the 3 fiducials occurs in each sub-aperture and that there are therefore 3 fiducial residuals to compute for each optimizer iteration.

6.1.2 Results

Figure 6.1 shows the original sub-aperture data where the region bounded by the black lines is the overlap area. The raw and aligned stitching results are shown in

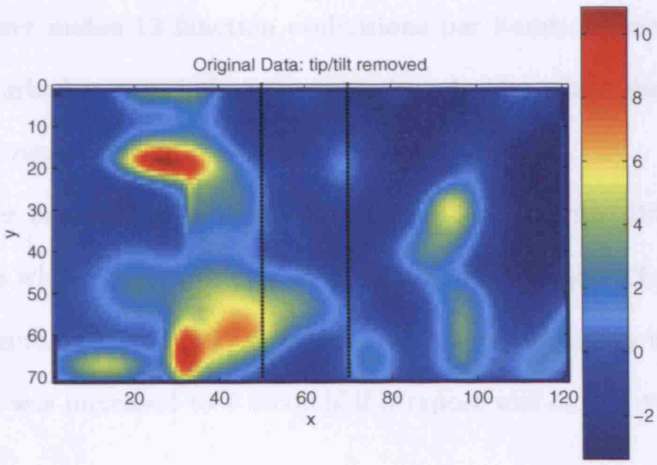


Figure 6.1: Simulated sub-apertures in their home positions. The overlap area between the two sub-apertures is the region bounded by the black lines.

figures 6.2 and 6.3 respectively, and the optimization statistics are given in table 6.1, and show the optimizer's progress in successively reducing the object function value

until the required tolerance was met.

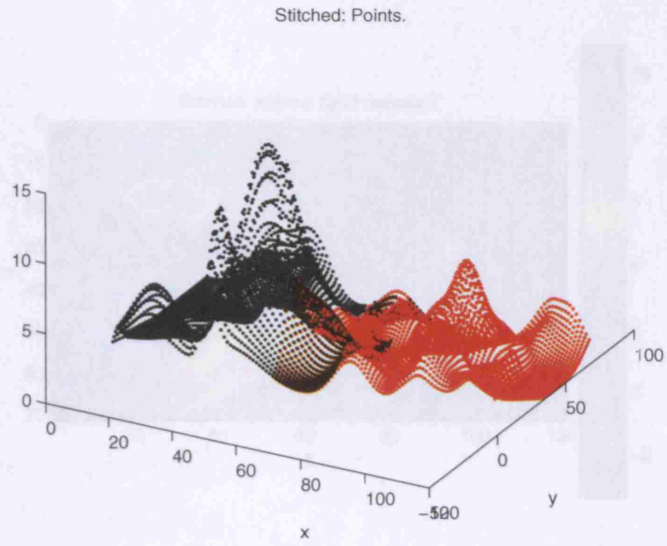


Figure 6.2: Raw stitching result for the two sub-apertures showing the discrete data points.

The optimization is in 12 variables – 2 sub-apertures, 6 variables for each – and the optimizer makes 12 function evaluations per iteration[1] where each design variable is perturbed in turn before the optimizer decides which direction in which to move for the next iteration.

As a further experiment, a series of runs was made with different values of `DiffMinChange` whilst keeping all other parameters constant. The results of this are shown in figure 6.6. The CPU time taken for the optimization in table 6.1 was 1 second which was increased to 4 seconds if a report was output to the console at each iteration[2].

[1]For n sub-apertures, the number of variables is $n \times 6$ and this corresponds to the number of function evaluations per iteration.

[2]The time taken is approximate because other tasks may be using the CPU in a multitasking environment.

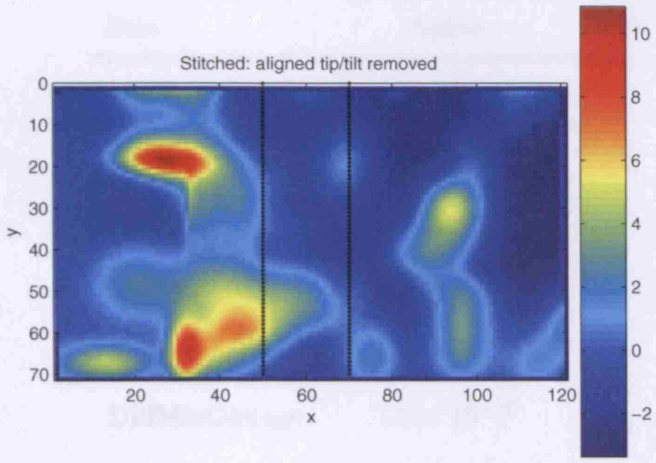


Figure 6.3: Stitching result for the two simulated sub-apertures.

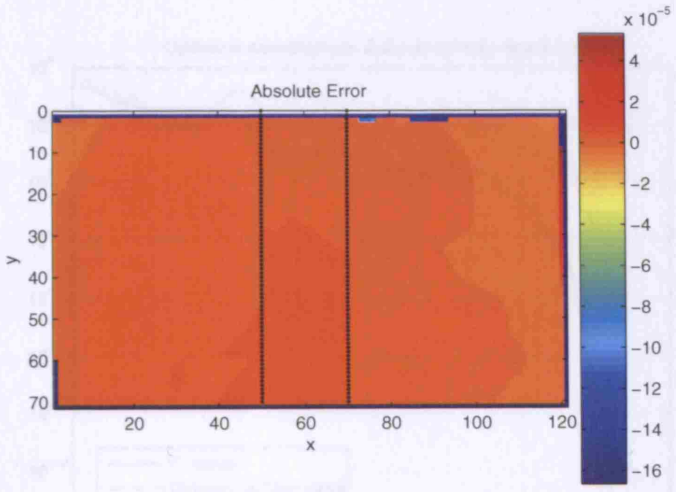


Figure 6.4: Error map for the stitching. This is computed as a point-by-point difference between the original data and the stitched data.

Stat	Value	
No. Iterations	73	
Object fun val (f)	5.2×10^{-12}	
Exit flag	1	
TolFun	1.0×10^{-10}	
TolX	1.0×10^{-11}	
DiffMinChange	1.0×10^{-16}	
PV Error	2.2×10^{-4}	units
RMS Error	5.9×10^{-6}	units
$\bar{r}$	1.3×10^{-6}	units

Table 6.1: Statistics for the stitching of the two simulated sub-apertures.

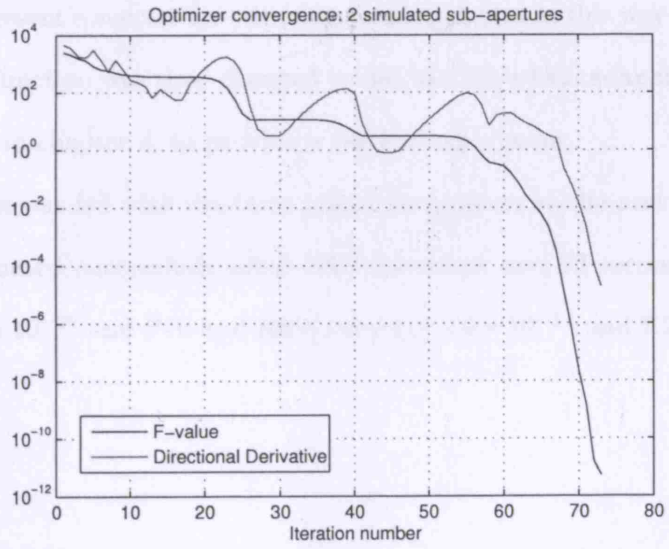


Figure 6.5: Plot showing the convergence of the object function and the magnitude of the directional derivative as a function of the iteration number.

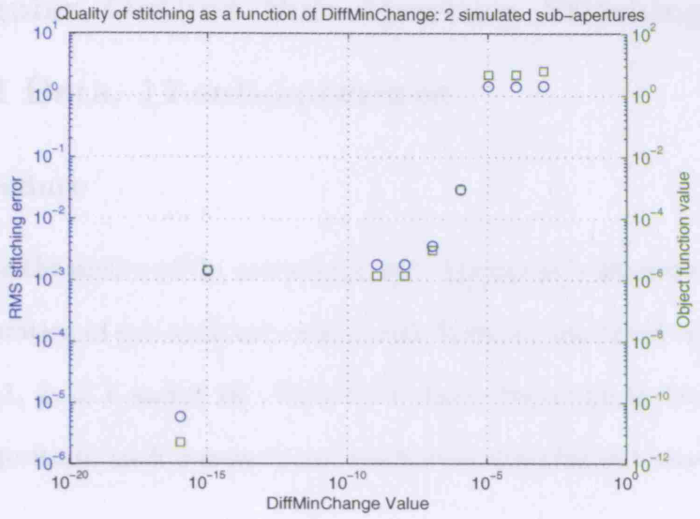


Figure 6.6: Results of varying the value of `DiffMinChange` in the optimizer options. Above values of 1×10^{-6}, the RMS errors and the object function values are comparable in size to the features in the data. Below this the stitching improves markedly until at `DiffMinChange` = 1×10^{-16}, where no further improvement was made – choosing 1×10^{-40} made no difference and suggests the limit had been met.

6.1.3 Discussion

The results represent a success for stitching simulated data in this way with no noise. The optimizer function was then changed to use Matlab's `fminsearch`, as described in section 4.5.2 in chapter 4, to provide a further experiment.

The function was fed with the same initial parameters as the previous optimization and terminated successfully after 2900 iterations and 73 seconds to give the f-value of 8.22×10^{-26} and P-V and RMS errors of 3.0×10^{-11} and 7.2×10^{-13} units respectively.

6.2 Irregular Outline Sub-Aperture Stitching: Simulated Data, 12 sub-apertures

6.2.1 Procedure

To demonstrate the ability of the software to stitch irregularly shaped objects, an L-shaped configuration of sub-apertures was chosen from the simulated set, comprising sub-apertures 1, 2, 5, 6 and 9–16. Figure 6.7 shows the original data set. Again to test the algorithm, each sub-aperture patch was transformed randomly before stitching.

6.2.2 Results

Figure 6.8 shows the stitched sub-apertures after optimization. The optimizer took

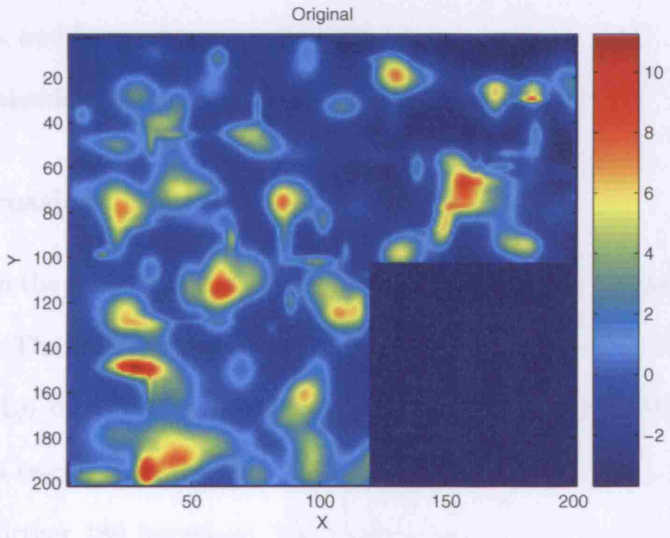


Figure 6.7: Original sub-aperture data for stitching 12 sub-apertures. The blue area at bottom right represents no data, and generally anywhere outside the aperture is assigned the value `NaN` in this work and thus is excluded in the evaluation of surface statistics..

much longer here, typically taking over 1000 iterations before terminating, and each iteration taking just less than 1 second. The error map of this stitching run is shown

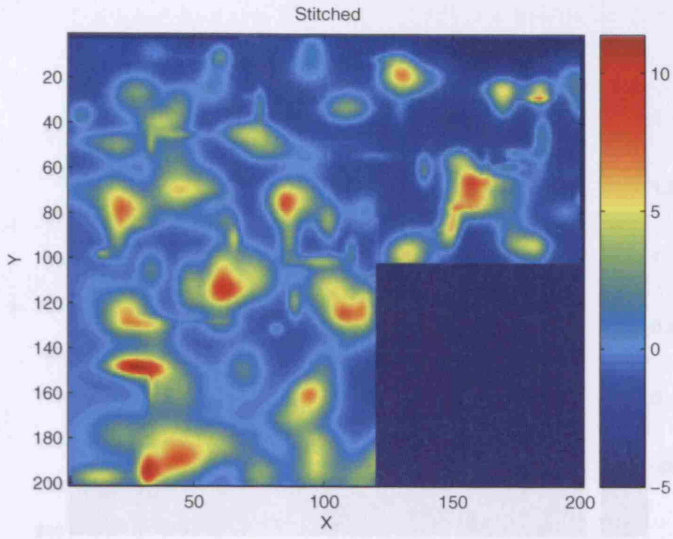


Figure 6.8: 12 stitched sub-apertures in L-shaped configuration.

in figure 6.9 The stitching here takes much longer due to the increased number of sub-apertures, and hence design variables.

The optimization statistics are given in table 6.2 on page 236

6.2.3 Discussion

It is clear from the table that the optimizer did not terminate because the tolerances were too low. The tolerances used can be decided on previous stitching runs based upon knowledge of the likely measurement and fiducial centroiding errors. The optimizer was re-run with the final optimized values for this run as the new seed and after a further 180 iterations, the f-value was reduced to 0.47, and the RMS error to 0.27. This may occur for several reasons:

- This is possible since the optimizer starts from $\mathbf{x}^{(0)}$ and proceeds in a slightly different initial direction from when the optimizer previously terminated, since no previous data is available.

- The optimizer may choose a different step size in the line search.

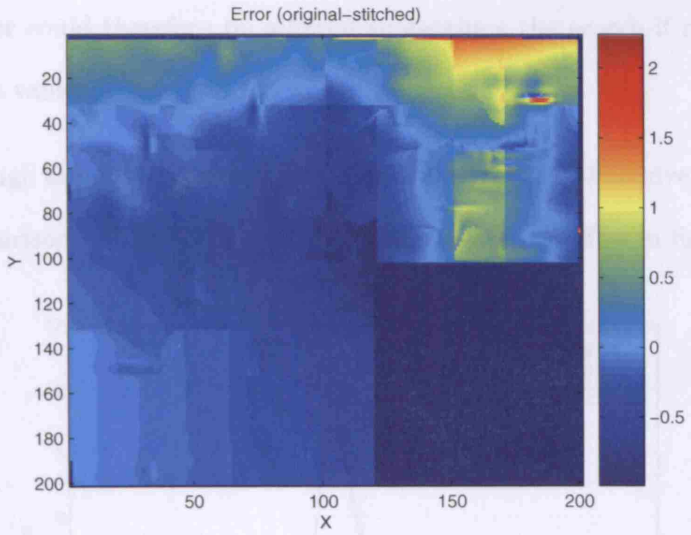


Figure 6.9: Error map for the 12 sub-aperture stitching. The errors are centred around zero except for in the top-right of the plot there they rise significantly.

Stat	Value	
No. Iterations	>1000	
Object fun val (f)	0.75	
Exit flag	-2	
TolFun	1×10^{-7}	
TolX	1×10^{-8}	
DiffMinChange	1×10^{-20}	
PV Error	2.95	units
RMS Error	0.34	units
$\bar{r}$	0.12	units

Table 6.2: Statistics for the stitching of the 12 simulated sub-apertures.

- The maximum number of iterations or function evaluations was reached - the optimizer could therefore be allowed to continue the search if re-started from previous values.

Sections through the data are plotted in figures 6.10 and 6.11 to give an alternative form of comparison. The RMS statistics for the error profiles in figures 6.10 and

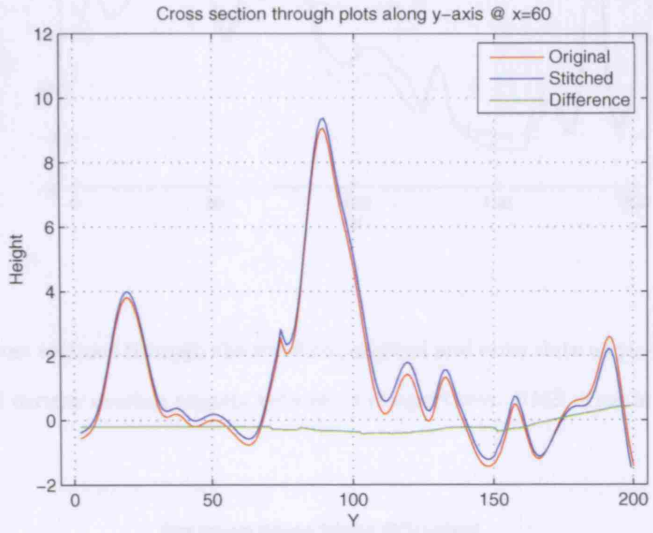


Figure 6.10: Cross sections through the stitched, original and error data at x=60 in figures 6.7– 6.9. All data at x=60 occupy overlap regions between sub-apertures. RMS error is 0.09.

6.11 are 0.09 nm and 0.14 nm respectively.

As a further result, the optimizer function was changed to `fminsearch.m`, using a simplex method, and started with the parameters found at the previous optimizer termination. The error map of this optimization is shown in figure 6.12 for comparison, with an RMS error of 0.14 units, 10000 iterations and 30 minutes of CPU time of which it is estimated that 50% of this time was taken up by outputting diagnostics information at each iteration.

Figure 6.13 shows the results of an experiment to attempt to reduce the errors further by running the optimizer for 5000 iterations at a time with an object function tolerance that is impossible to achieve. The f-values are plotted against the

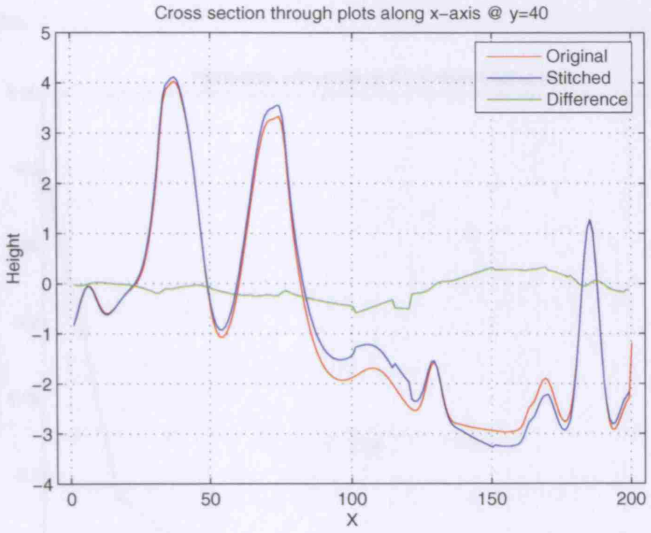


Figure 6.11: Cross sections through the stitched, original and error data at y=40 in figures 6.7– 6.9. All data at y=40 occupy overlap regions between sub-apertures. RMS error is 0.14.

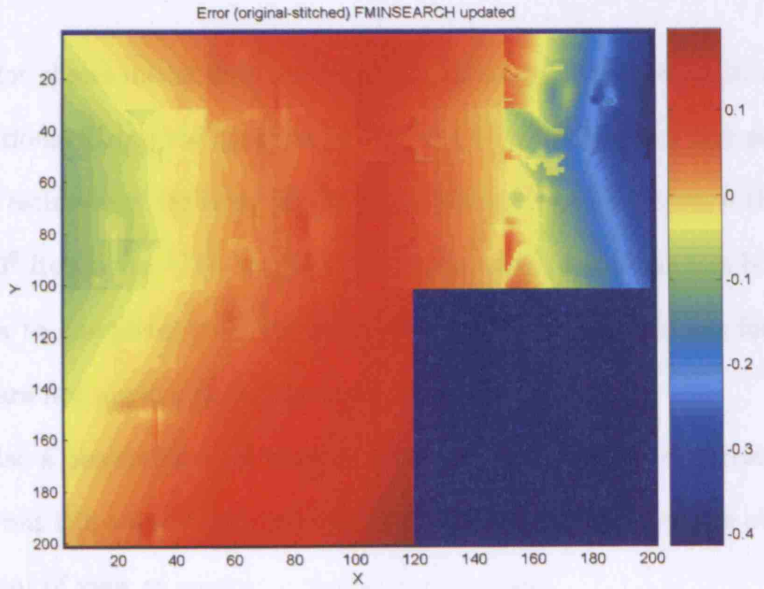


Figure 6.12: Revised error map for stitching the 12 simulated sub-apertures. A different optimizer function was invoked using a simplex search method. The RMS error is 0.11 units as characterized by the bulk colour of the plot.

cumulative iteration number is an effort to establish the trend for the convergence with ideal data.

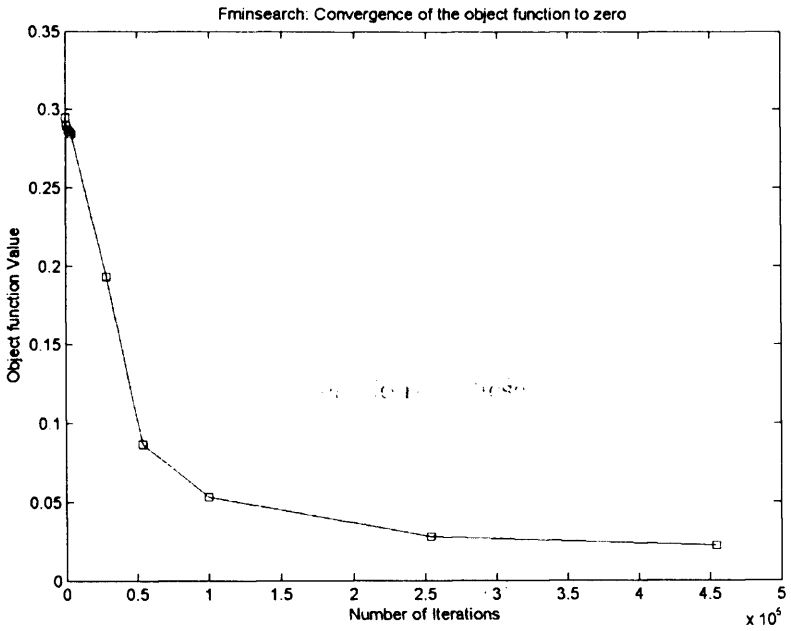


Figure 6.13: Convergence of `fminsearch` for the multi-aperture stitching in this section.

The plot shows initial high convergence with a marked change in gradient after 5000 iterations. Using the gradient of the line between the last two points gives an optimistic estimate of the time required to reach an f-value of zero, if this is possible, of 1.2×10^6 iterations. This would take approximately 300 minutes based upon the time taken to complete 4.5×10^5 iterations. These arguments are for `fminsearch` only and are not applicable to the more advanced optimizers.

It is also a possibility that there is some truncation error occurring in the computation, but this would require a detailed study to be made of the object function from a point of view of computer numerical analysis.

6.2.4 Conclusions

Stitching in this case has required a much higher number of iterations to reach an answer that approaches acceptability. This is expected as times for simple opti-

mizations such as the simplex method usually scale as n^2, where n is the number of variables. Here, `fminunc` and `fmincon` are more sophisticated and scale at less than this, but there is also a lot of housekeeping to be done that certainly scales as n^2.

6.3 Real Data

A part was sought that could be tested full-aperture and also via a series of sub-aperture tests to provide a yardstick for the stitching on real parts with all the sources of noise present. UCL has a limited stock of good parts from the former Grubb-Parsons company consisting of various test plates and lenses[3]. Many of the components were too small, but a good test plate was selected from a selection in the 200 mm diameter range.

6.3.1 Fizeau Interferometers and Optical Testing

Following on from the outlines of Fizeau interferometry given in chapters 1 and 2, some of the salient facts pertaining to the actual measurement of a surface are addressed that are important to consider for stitching.

In a Fizeau interferometer the datum and test wavefronts are generated by division of amplitude at the reference surface of the transmission element. The datum wavefront remains inside the interferometer where it interferes with the returned test wavefront. Any modifications to the test wavefront are encoded in the interference pattern produced.

A plane reference surface called a transmission flat is used for testing plano surfaces and the maximum size of the part that can be tested is the same size as the transmission flat. The scheme is illustrated in figure 2.9 on page 108.

[3]The Grubb-Parsons company completed the optics for the William Herschel Telescope in 1988, following this the parent company NEI Parsons decided that the future of the company lay outside optics. The company's optical assets were then sold off and purchased by UCL for the sum of £1500.00.

Transmission spheres are used to test spherical objects and generate spherical wavefronts which converge down to a point. The part to be tested is placed in front of the interferometer such that the tangents at all points on the test surface match the tangents of the wavefront incident upon it, this then gives a null test where the elements are confocal as illustrated in figure 6.14.

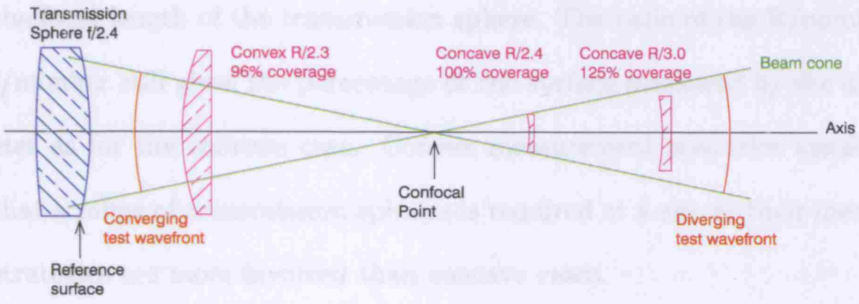


Figure 6.14: Confocal testing of spherical parts.

Transmission flats are specified by their diameter whereas transmission spheres are specified using their focal ratio and diameter. The focal ratio, or speed, is simply the traditional photographer's f/#, i.e., the ratio of the focal length to the clear aperture. The analogous quantity for a surface is the R/# defined as the ratio of the radius of curvature to the clear aperture as used by Zygo [116].

Several important points require consideration when using transmission spheres

1. **Transmission sphere faster than the concave SUT**: If the transmission sphere is faster than the SUT, where the SUT is concave, then the whole of the surface under test can be measured in one shot.

2. **Transmission sphere slower than the concave SUT**: A single full-aperture test is not possible and a sub-aperture test must be devised. If the interferometer beam does not cross the edge of the SUT then the ratio of the R/number to the f/number gives the percentage of the surface measured by the interferometer.

3. **Convex SUT**: Full-aperture coverage of convex surfaces is difficult if they are of a diameter more than a small fraction of the interferometer transmission sphere diameter. The confocal point is virtual and located behind the surface under test in the direction of the outgoing test beam. It is not possible to measure any portion of a convex surface if its radius of curvature is longer than the focal length of the transmission sphere. The ratio of the R/number to the f/number still gives the percentage of the surface measured by the interferometer as for the concave case. Convex measurement scenarios usually dictate that a range of transmission spheres is required at a site as their measurement strategies are more involved than concave cases.

4. **Slow and concave SUT**: If space is an issue it is possible to obtain diverging transmission spheres where the confocal point is virtual. Their drawback is that they cannot be used to test convex surfaces and that the radius of curvature of the SUT must be longer than the focal length of the transmission sphere since the focus is virtual.

Factors (2) and (3) are where sub-aperture testing can be used to enable the whole of a surface to be measured that cannot be measured in (1) above, and the case where a full-aperture measurement cannot be made of a convex surface. Diverging elements in (4) are used only in special cases where the distance between the SUT and the interferometer needs to be minimized.

Diverging elements may also be used to test parts with very large radii of curvature. For example, ESO's OWL had a spherical primary with a radius of curvature of 100 m. Testing with a converging transmission sphere would require an optical path length of over 100 m which would be subject to turbulence, stratification and vibration.

Instead, a diverging transmission element could be used to provide a wavefront with the required radius of curvature close to the interferometer and here we would

be in the situation of (2) in the list above where sub-aperture testing could be used. The swing-arm profilometer could be used to servo the interferometer in this instance as the interferometer could be made to have its focus coincident with the centre of curvature of the SUT at all times. To obtain the sub-aperture measurements the part could be rotated and the arm swung to enable each point on the surface to be measured.

6.3.2 Sub-aperture layout design

The parts available for testing the stitching algorithm were flats and concave spheres. The simplest arrangement of sub-apertures for a concave sphere is a series of concentric sub-apertures as exemplified in figure 6.15. Such arrangements have been demonstrated by Fleig et al, with the QED stitching technology [65].

6.4 GP_11

6.4.1 Procedure

A concave spherical test plate, referred to hereafter as GP_11, was selected from the UCL stock of Grubb-Parsons plates to test the stitching algorithm. Testing took place in the confocal arrangement.

The part was a concave sphere in low expansion borosilicate glass with a spherometer-measured radius of curvature of 283.74 mm and a diameter of 205.0 mm, giving $R/1.4$. Unlike some of the test plates the surface of GP_11 was free from scratches and other defects, and was known to have a form error of less than $\lambda/4$ P-V at 633 nm [117].

To give a benchmark for the stitching, the UCL Specac $f/0.9$ transmission sphere was used on the Wyko 6000 to test the part full-aperture, giving a theoretical coverage of 156%. The sphere was calibrated with a $\lambda/10$ P-V reference sphere.

The scheme devised to test the part used 4 sub-apertures to give 100% surface

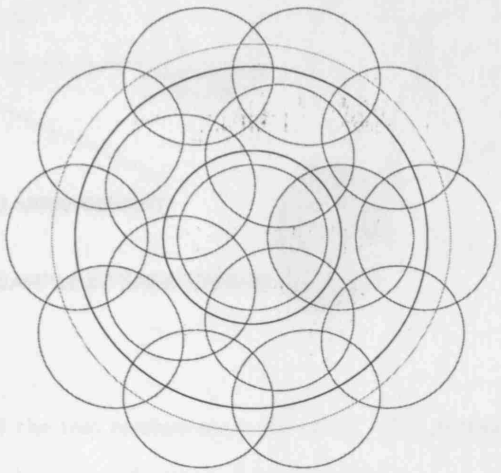


Figure 6.15: For axially symmetric parts, simple arrangements of concentric sub-apertures are most easily generated as shown here. The green circle represents the outline of the surface under test, the red lines are the extents of each sub-aperture measurement and the black lines are the concentric pitch circles for the sub-aperture measurements. The symmetry makes composing the optimizer seed transformation parameters easier. If the part is plane then moving between sub-apertures is achieved by translating the part relative to the interferometer about a plane perpendicular to the optical axis. If the part is spherical then the part is rotated about its (and the interferometer's) centre of curvature. The situation for aspherics is complicated by the fact that a relative translation had to be made along the optical axis to account for the change in curvature of the part, and that the size of this translation needs to be tracked very precisely for each sub-aperture measurement.

coverage with significant overlap at the centre of the part. The 4 sub-apertures were arranged symmetrically about the part's axis according to the schematic shown in figure 6.16. Further to the specification of the sub-aperture arrangement, the

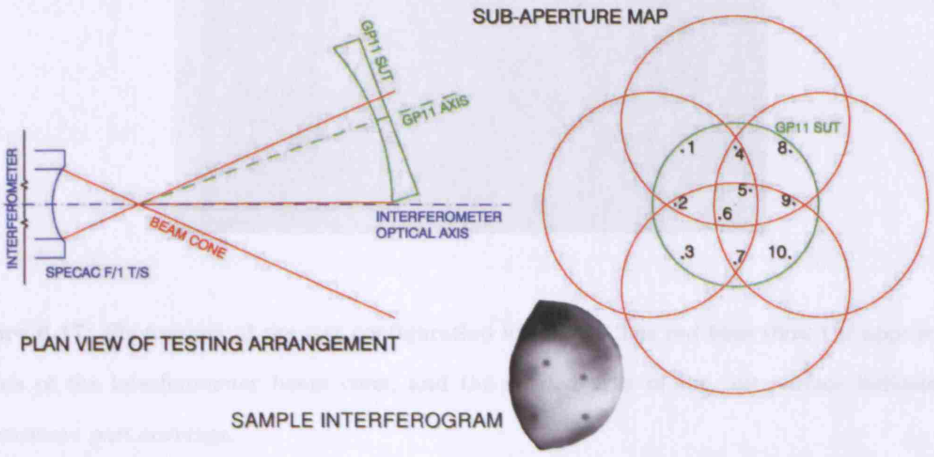


Figure 6.16: Schematic of the test configuration for GP11. The sub-aperture map on the right was scaled and attached to the part to facilitate initial alignment.

part was fiducialized by hand, shown and numbered in figure 6.16. Fiducials were approximately 5 mm in diameter.

The part was placed in a high-quality manually operated tip/tilt and focus stage. Initial positioning was accomplished with the aid of a computer generated plot at actual-size showing the extents of the interferometer beam at the 4 sub-aperture configurations. This plot was attached to the surface under test and the beam extents were made to coincide with the plotted outlines by moving the part and mounting relative to the interferometer.

Each measurement was therefore made off-axis with respect to the part's axis of symmetry which was inclined to that of the interferometer. The arrangement was such that the tangents of the test wavefronts were parallel to the tangents of the surface under test at all times, to give an off-axis null measurement at the centre of curvature[4]. A photograph of the whole set-up is shown in figure 6.17 showing the

[4]A traditional off-axis test using a return sphere introduces other aberrations into the test.

approximate extents of the interferometer beam.

Figure 6.17: Photograph of the test configuration for GP11. The red lines show the approximate extents of the interferometer beam cone, and the shaded area of the test surface indicates the approximate part coverage.

Following initial rough alignment the part was carefully aligned to interfere with the beam from the interferometer. At this point the fringes were nulled by removal of tip/tilt and defocus using the precision mount and the surface form was measured by the interferometer. The symmetry of the sub-aperture configuration meant that the part had only to be rotated by 90° in-situ and nulled again between each measurement. After the 4 sub-aperture measurements were taken, the part was measured full-aperture with it's axis coincident with that of the interferometer.

The initial values for the optimizer were simply the values of the rotation angles about the z-axis used: 0° , 90° , 180° and 270° . All other seeds were set to zero.

6.4.2 Results

After testing, the surface deviation maps were imported into the stitching software where the fiducials were centroided and the stitching software was run on the measurements. To illustrate the processes of chapter 4, the phasemaps at each stage of the processing are included here as this will aid further development work on the algorithm.

Raw phase measurements are shown in figure 6.18 and show some variation in the features of the surface of GP_11 which will all contribute to the stitching errors. However, some of the perceived variation is caused by the tip/tilt terms for the individual deviation maps being slightly different. The full-aperture map shows the part's P-V form error to be approximately 140 nm and the Matlab software computed the RMS form error to be 32 nm, in good agreement with the Wyko software analysis at the time of capture.

Following the procedure in section 4.1.1, the corrections were made for the transmission sphere used and the location of the optical axis. Now the edge form of the sub-aperture measurements appears circular and each map is correctly scaled and these corrected but unstitched deviation maps are shown in figure 6.19. Note that the plotted data have had the missing fiducial points interpolated over. At this stage in the software processing, the actual points stored in the arrays still have the data points missing.

After stitching the new array containing the stitched measurement result was generated from the four transformed sub-apertures shown in figure 6.20. The sub-apertures as plotted in their stitched states at this point are subject to a degree of tip/tilt error, i.e., each sub-aperture measurement has an individual degree of tip/tilt that is distinct from the 4 when considered as a whole. This is only corrected fully in the stitched array data as the plots of figure 6.20 are only used as a diagnostic aid in case of user error.

Table 6.3 outlines the statistics for the optimization, showing the optimistic tolerance values and the lowest f-value obtained.

The unfiltered stitched deviation map is shown in figure 6.21 after tip/tilt correction and clearly shows the spatial frequencies introduced by the stitching process. For comparison, the full-aperture map is shown in the lower half of figure 6.21.

Filtering was carried out on the data as per section 4.8 using a simple 10×10

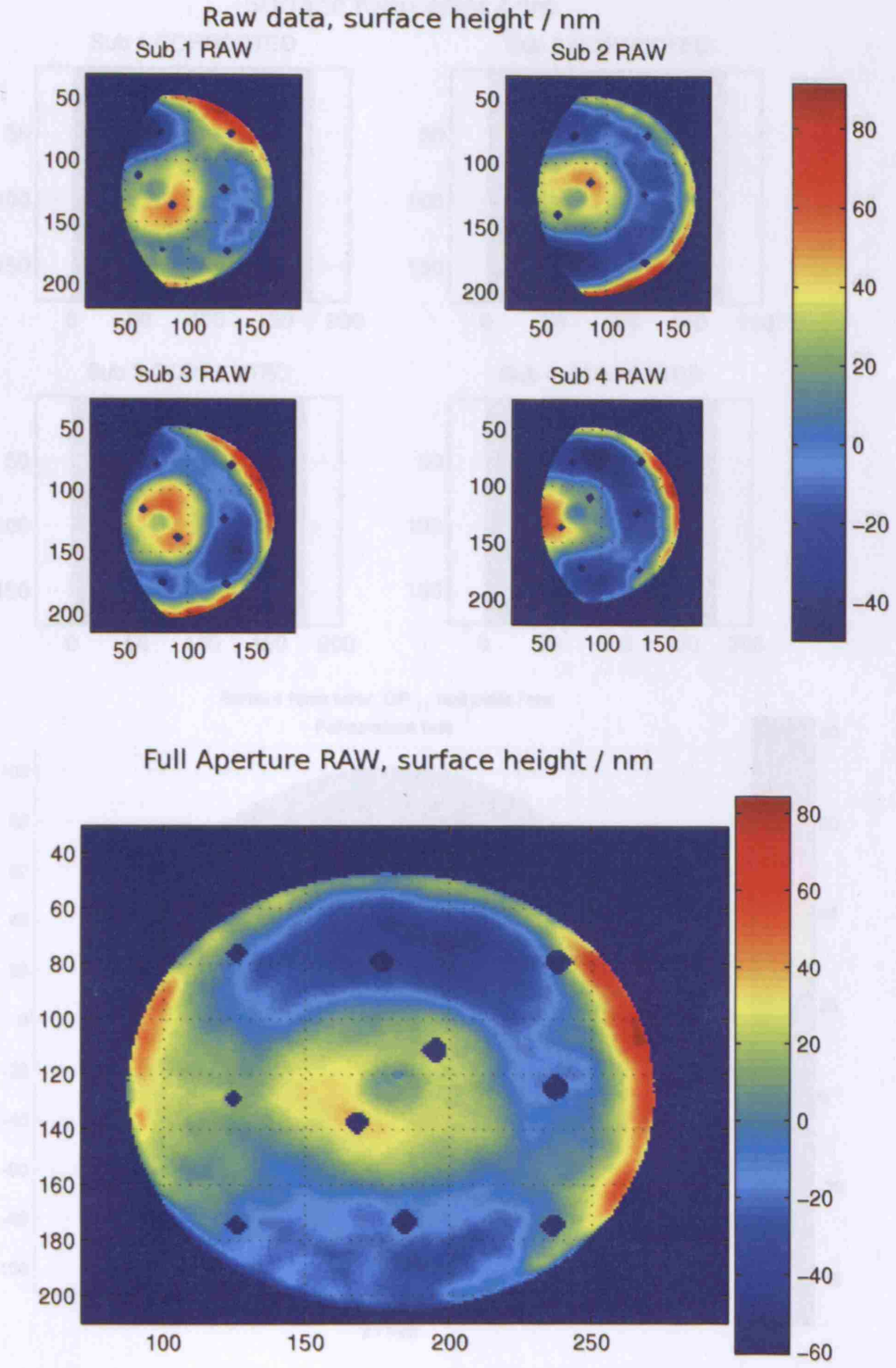


Figure 6.18: Upper: Raw sub-aperture data from Wyko 6000. Note the incorrect aspect ratio. Lower: Full-aperture data showing the fiducials.

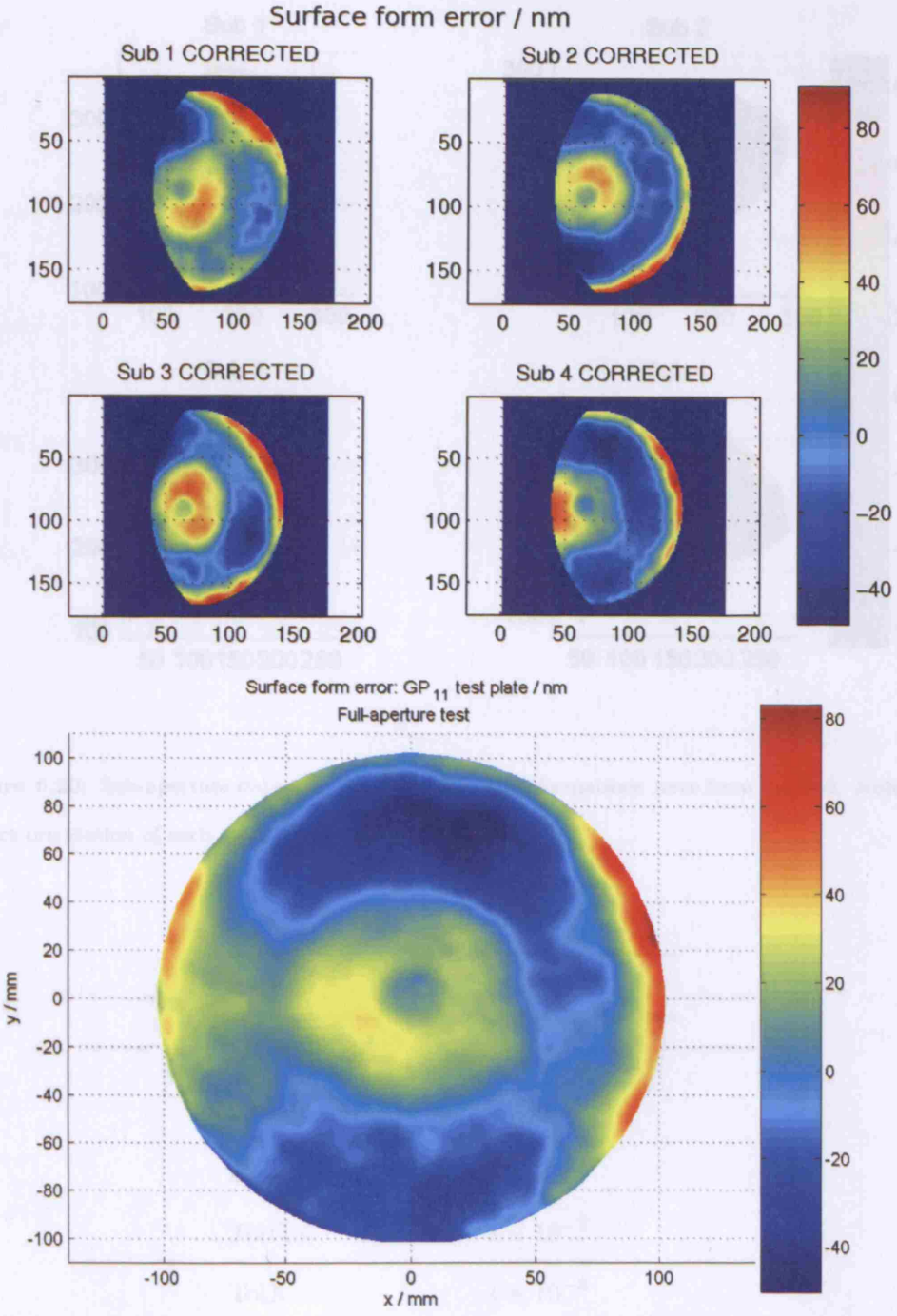


Figure 6.19: Upper: Raw sub-aperture data corrected for the coordinate system and aspect ratio. Lower: Corrected full-aperture data. Note that at this stage, the missing data points at the fiducial locations have been interpolated over by the plotting routine.

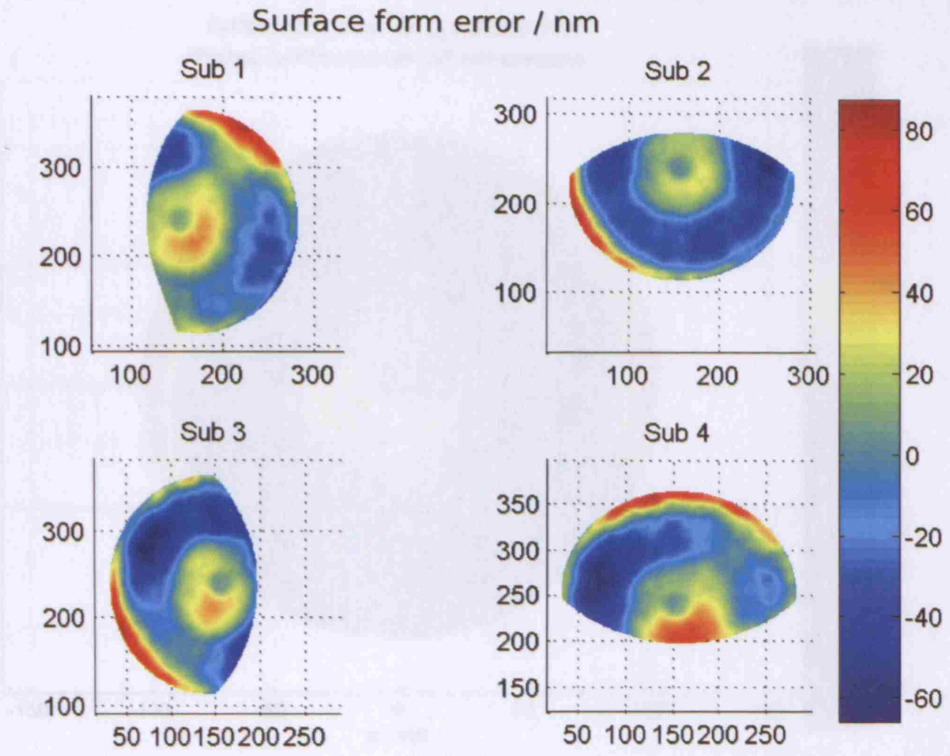


Figure 6.20: Sub-aperture data after the stitching transformations have been applied. Note the correct orientation of each map.

Stat	Value
No. Iterations	508
Object fun val (f)	6428
Exit flag	-2
TolFun	1×10^{-7}
TolX	1×10^{-8}
DiffMinChange	1×10^{-20}
$\bar{r}$	15 nm

Table 6.3: Statistics for the stitching measurement of GP_11 in 24 variables. CPU time used was 60 seconds. Note the very optimistic tolerances.

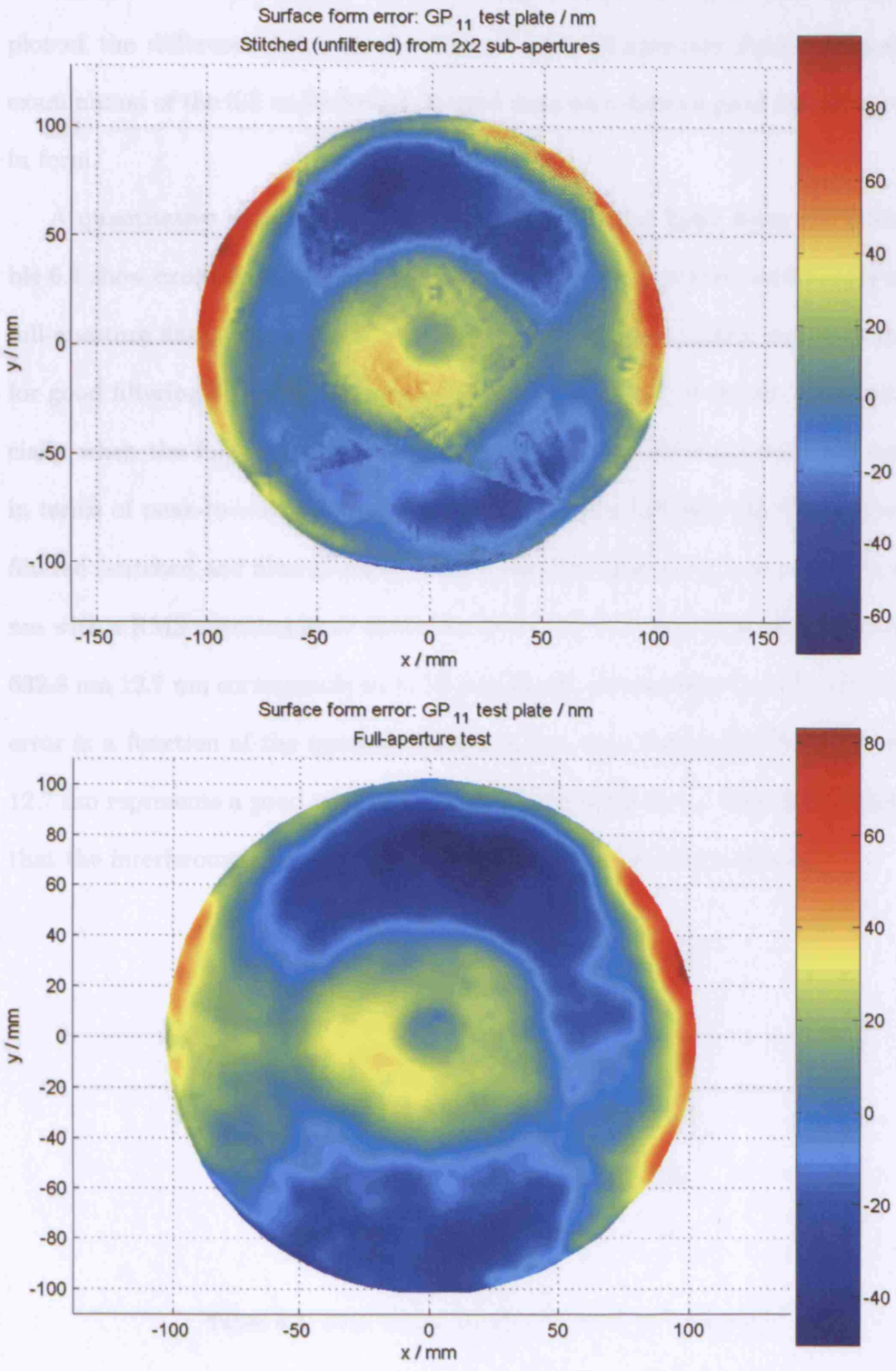


Figure 6.21: Upper: Stitched data. Note the geometrical patterns due to overlapping grids of data. Lower: Full-aperture for comparison.

averaging filter. The filtered deviation map is shown in figure 6.22 along with a plot of the difference between the filtered and full-aperture data. A qualitative examination of the full and stitched-filtered data sets shows a good overall agreement in form.

A quantitative examination is more revealing. The basic form statistics in table 6.4 show excellent agreement between the filtered (stitched and filtered) and the full-aperture data. The matching between these numbers clearly indicates the need for good filtering to remove the noise from stitching and to recover the form, especially when the full-aperture data is compared to the stitched and unfiltered data in terms of peak-to-valley form error. The difference between the full-aperture and filtered (stitched and filtered data) reveals that the peak error was approximately 80 nm with a RMS stitching error (point for point) of 12.7 nm. At a laser wavelength of 632.8 nm 12.7 nm corresponds to $\lambda/50$ and 80 nm corresponds to $\lambda/7$. The stitching error is a function of the quality of the original data and many other factors, but 12.7 nm represents a good result for the RMS stitching error. Especially considering that the interferometer was calibrated with a $\lambda/10$ PV return sphere.

Form error	PV	RMS	
Full-aperture	144.6	22.7	nm
Stitched	170.4	27.6	nm
Filtered	148.1	25.2	nm
Full-filtered	86.4	12.7	nm

Table 6.4: Basic form statistics for the GP_11 test plate.

As a further measure of the stitching error, cross sections through the stitched and full-aperture data were computed and are plotted in figure 6.23 with the corresponding errors in figure 6.24. The profiles through Y are in very good agreement

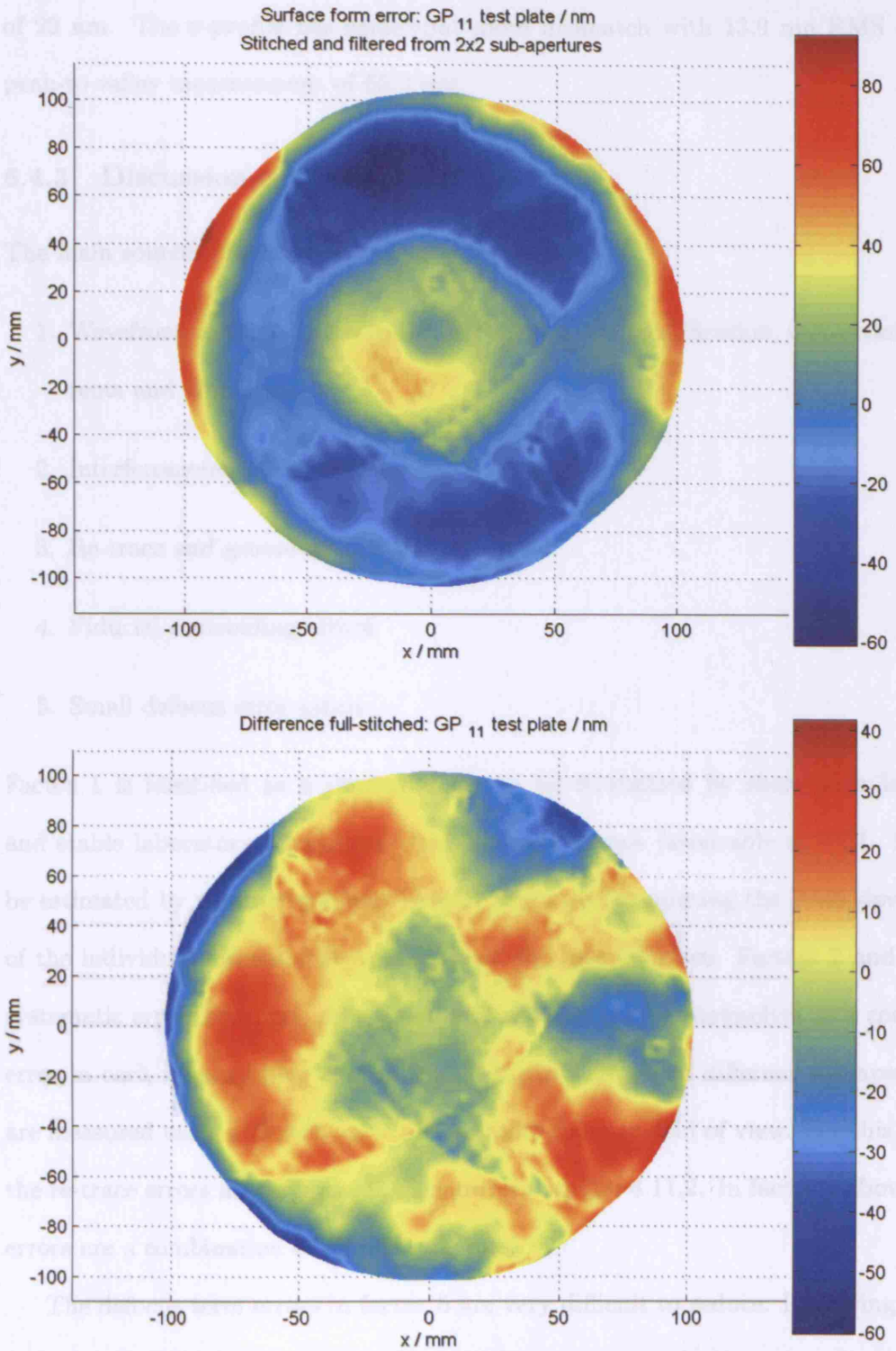


Figure 6.22: Upper: Stitched and filtered data. Lower: stitching error map.

and the RMS error in the y-profile is 6.3 nm with a peak-to-valley measurement of 29 nm. The x-profile has somewhat more mismatch with 13.9 nm RMS and a peak-to-valley measurement of 62.2 nm.

6.4.3 Discussion and Error Estimation

The main sources of error are identified as:

1. Wavefront distortions due to air turbulence, air stratification, convection currents and vibrations.
2. Interferometer calibration errors.
3. Re-trace and geometric distortion errors.
4. Fiducial centroiding errors.
5. Small defocus error terms.

Factor 1 is identified as a random error to be minimized by short path lengths and stable laboratory conditions which are not always favourable at UCL. It can be estimated by measuring a part several times and computing the RMS deviation of the individual measurement points from the mean surface. Factors 2 and 3 are systematic errors and such calibration errors will manifest themselves as a constant error in each measurement and represent a problem where different sub-apertures are measured using different parts of the interferometer field of view. For this work, the re-trace errors are very small, as shown in section 4.11.2. In factor 4 above, the errors are a combination of all other factors.

The defocus term errors in factor 5 are very difficult to reduce. Removing aberration terms from interferometry data involves computing least-squares fits to a set of orthogonal basis functions - in this case, the Zernike polynomials defined upon the unit circle – a circle in $\mathbb{R}^2$ with unit radius. We thus assume that the wavefront

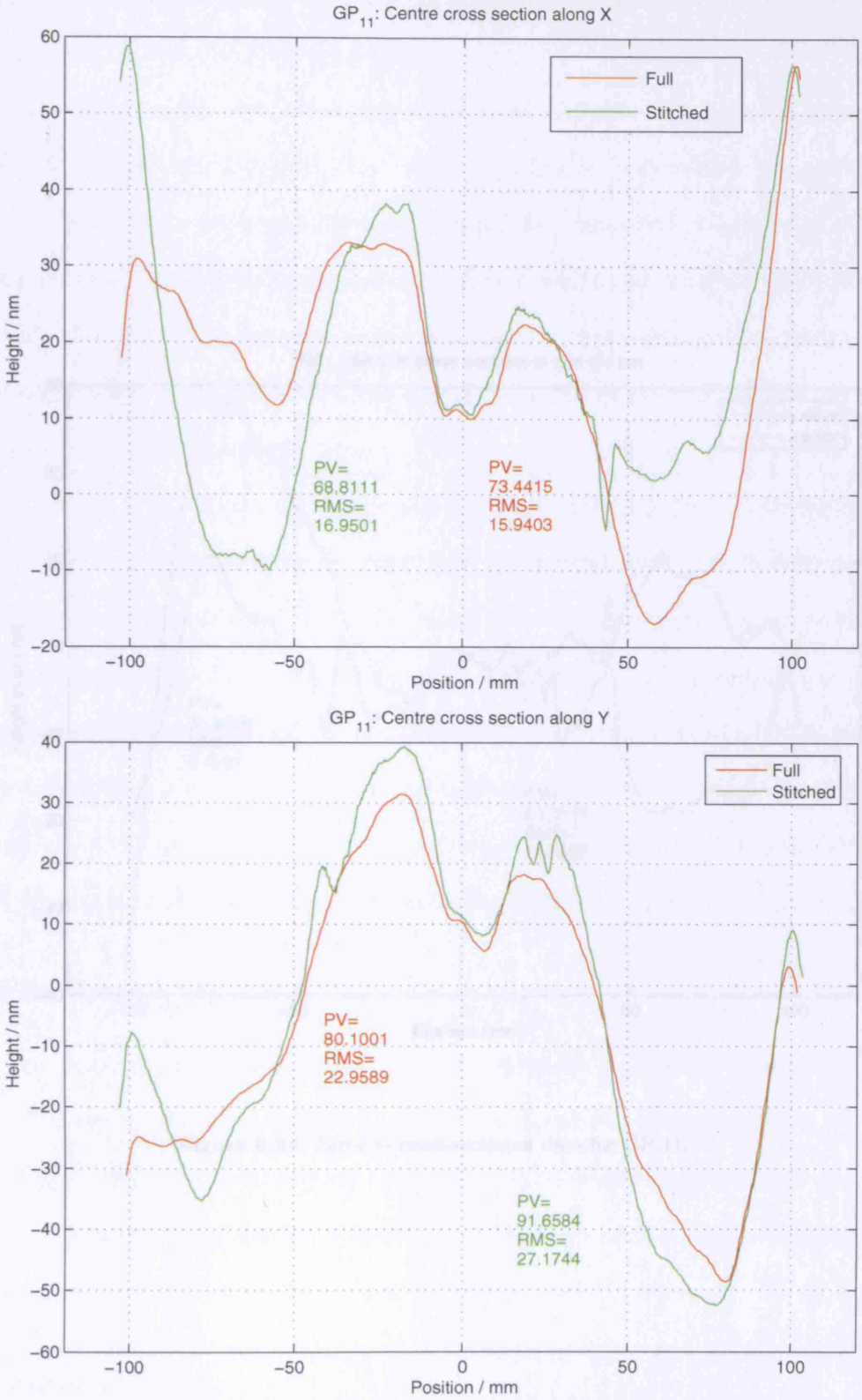


Figure 6.23: Cross sections through stitched and full-aperture data in X and Y for GP_11.

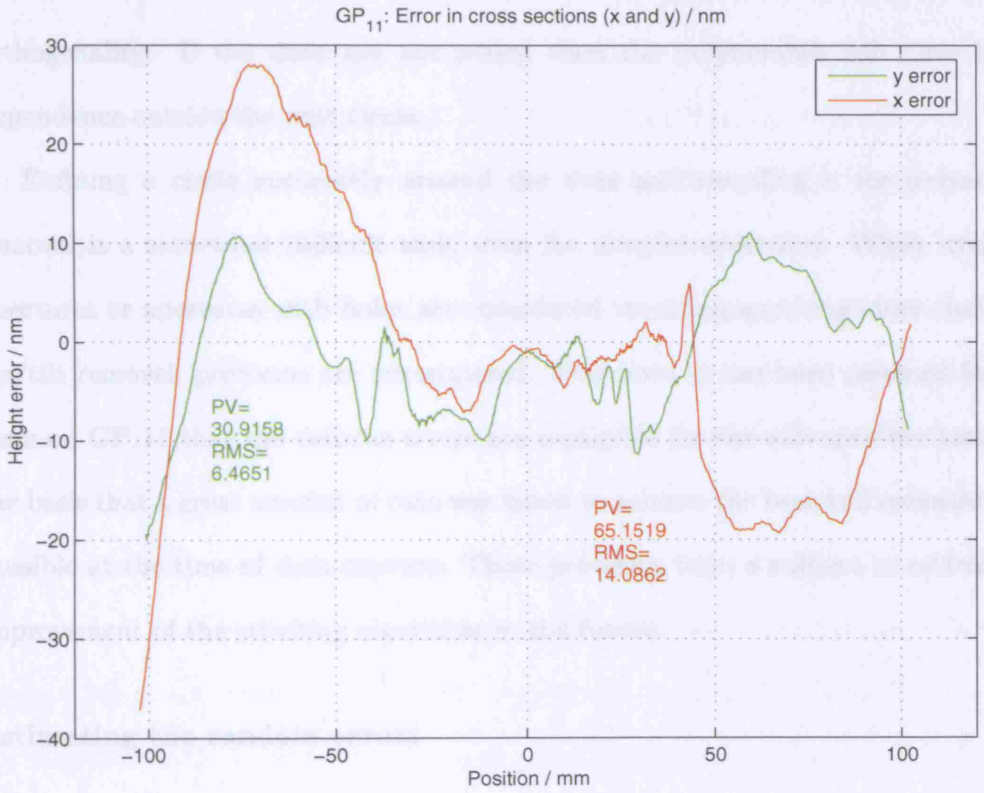


Figure 6.24: Error in cross-sectional data for GP_11.

can be expressed as an infinite series of orthogonal functions – defocus is one such polynomial in the Zernike system.

The first two Zernike polynomials define tip and tilt and these are orthogonal over $\mathbb{R}^3$. However, after this, the remaining Zernike polynomials are orthogonal within the unit circle only, although they can be computed outside of it. The x- and y-coordinates of the phase data require rescaling so that the data just lies within the unit circle, then the aberration Zernikes can be removed whilst ensuring orthogonality. If the data are not scaled then the polynomials will have linear dependence outside the unit circle.

Defining a circle accurately around the data and rescaling it for polynomial removal is a somewhat difficult task, even for circular apertures. When irregular apertures or apertures with holes are considered requiring anything more that just tip/tilt removal, problems are encountered. Therefore, it has been assumed for the tests on GP_11 that the defocus terms are negligible for the sub-aperture tests, on the basis that a great amount of care was taken to achieve the best null measurement possible at the time of data capture. These problems form a subject to address for improvement of the stitching algorithm in the future.

Estimating the random errors

A part was measured 6 times in the same configuration and used to estimate the random errors. The mean surface form was computed and used to compute the point by point errors in the individual phasemaps. The RMS error was then computed as 3.7 nm and can be used to account for the random errors due to turbulence and vibration.

Systematic Errors

Interferometers are usually calibrated using a known spherical or flat reference object. The form is measured several times and an average is computed that is sub-

tracted from all measurements subsequently made. Most reference objects are certified to have a form error $< \lambda/20$ PV at 633 nm which corresponds to 32 nm. However, a more precise method than is available at UCL is required to fully characterise the interferometer's reference optics on the 1 nm level. At present, even averaging several measurements of a calibration object still results in a data set containing terms from both the transmission object and the calibration object, i.e. the reference wavefront contains features correlated to the calibration surface.

Parks [118] has suggested a calibration method for transmission spheres that could be used to improve this work, based upon measuring different parts of a good sphere several times. The sphere is placed confocal with the interferometer's transmission sphere and measured before being rotated. The procedure is repeated to give n-measurements and as they are uncorrelated (assuming different parts of the sphere were measured each time) then they will average out to zero but the systematic errors in the transmission sphere will add coherently and thus be very well estimated by the average of the n measurements.

Commercial parts are now available including a 25 mm diameter silicon nitride sphere and kinematic mount – the advantage of the silicon nitride is that it is hard, opaque[5], does not rust or tarnish and is spherical to 120nm P-V. Such a test was not available at UCL, but this is a recommendation for future work.

The re-trace errors are minimized when testing good spheres at null and become a real problem with aspherics where the outgoing and incoming rays proceed along different paths after reflection. This is also true with slower transmission spheres where the path lengths are longer. The geometric errors are particularly important for faster optics and very extensive calibration is required to reduce them. Inevitably, any imaging device will introduce geometric distortions of the data. One method to calibrate these out is the calibrate the imaging system with a reference image.

[5]There is no reflection from the rear concave surface.

Both the related errors mentioned in this paragraph require further study to reduce them.

The various errors can be used to estimate a target value for the object function of a given optimization, and is illustrated for the GP_11 part by considering the estimated error for a fiducial position.

There are 28 fiducial pairs in GP_11 and the results in chapter 2 indicate an x-y fiducial centroiding error of 2 pixels in each dimension, and a z-error of 3 nm. Let us also use the figure of 3 nm as a further random error from any interferometer measurement. For GP_11 a pixel on the interferometer corresponds to approximately 1.3 mm in the results of this section giving 2.6 mm for the x- and y-errors. We use mm for x and y, and nm for z to be consistent with the method used to compute the object function, and combining in quadrature to give the worst case gives:

$$\frac{r}{2} = \sqrt{2.6^2 + 2.6^2 + (3+3)^2} = 7.0 \tag{6.2}$$

For the worst case, 2 fiducials will be involved in a pair and will each have this error so we can conclude that $r = 14.0$. This compares with $\bar{r} = 15$ for the stitching result on GP_11. Comparing the estimated f-value with that obtained we have, from equation 4.7, $f_{exp} = 14.0^2 \times 28 = 5488$ and $f = 6428$ respectively, a favourable comparison, where 28 is the number of fiducial pairs and 14 units represents the worst error between two fiducials. In fact, we expect the real f-value to be larger since geometric distortion has not been fully accounted for in the interferometer optics. The value obtained is also in good agreement with the RMS stitching error computed as the difference between the full-aperture and stitched measurement shown in table 6.4.

The error analysis above suggests that the stitching procedure is working and finding an answer close to the best attainable under the circumstances of the test. The results indicate that the measurement errors have to be reduced to improve the stitching. Even when testing full-aperture, the precise calibration of the refer-

ence surface is a critical step in measurement of parts with form errors of a similar magnitude to those of the reference element.

A reduction in the fiducial centroiding errors and a good calibration of the reference optics is thus the first place to start in improving the stitching. The difficult subject of geometric distortion can then be addressed to further tie down the errors. No doubt a further reduction will be gained by testing in a more stable environment than UCL. Even on a vibration isolated bench, London underground trains provide a considerable source of low-frequency vibration as there are two lines less than 60 m from the metrology lab.

6.5 Fused Silica Flat

6.5.1 Procedure

A fused silica flat measuring 305 × 155 mm in size was selected for a larger stitching run using 2×3 sub-apertures. The flat was originally used to perform some homogeneity tests on a large prism for a spectrograph constructed by UCL and was known to be flat to $\lambda/4$ PV on one side and $\lambda/10$ P-V on the other at at 632.8 nm [117]. The flat's vertical dimension was just too large to be fully accommodated by the interferometers's 152 mm diameter aperture therefore the surface was divided up into 2×3 sub-aperture areas. The part was tested using the Wyko 6000 with a 150 mm diameter $\lambda/20$ P-V reference surface. Figure 6.26 contains a photograph of the test arrangements and the manually applied fiducials. Again, a full-size plot was made of the surface and the interferometer beam positions to aid alignment as shown in figure 6.25.

The measurement problem was made more difficult by the fact the two sides of the part were parallel and a second set of fringes was visible in the sub-aperture interferograms [119]. This can be dealt with in modern systems but not using the Wyko 6000. The rear-surface reflection was eliminated by oiling the fused silica flat

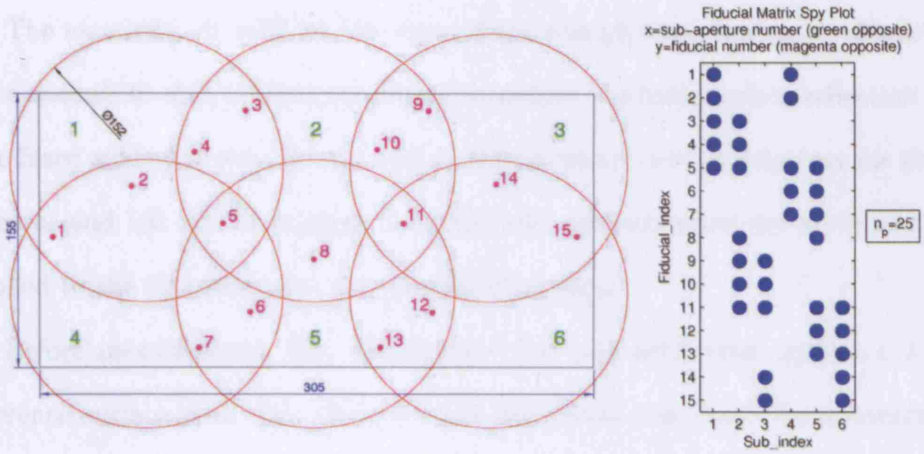


Figure 6.25: Sub-aperture layout for the fused silica flat and fiducial matrix spy plot.

Figure 6.26: Photo of the fused silica flat testing arrangement. The flat is resting on an aluminium bar on the wooden frame and is oiled to the large piece of glass whose rear side had been sandblasted to avoid further reflections.

to a high-quality piece of 1" thick plate glass whose rear surface had previously been shotblasted.

The matching oil used for the experiment was glycerol, having a refractive index close enough to that of the two glasses to reduce the back surface reflection enough. The fused silica flat was mounted in a custom mount with a drip tray for the excess glycerol and left for three days to settle into position after the glycerol had been applied to the interface with a hypodermic syringe.

Before measurement, the transmission flat was calibrated against a $\lambda/20$ P-V reference surface such that the reference wavefront error could be subtracted from each measurement. Sub-apertures 1–3 were first measured by simply translating the part mounting. Sub-apertures 4-6 were measured by raising the part mounting and then performing a similar translation. A second set of measurements of the part was made 6 days after the first to establish a measure of the repeatability of the stitching.

6.5.2 Results

A set of corrected sub-apertures is shown in figure 6.27 in order to compare the sub-apertures before $\mathbf{S}'$ in equation 3.26 on page 154 is formed. Figure 6.28 shows the results of the two stitching runs made on the fused-silica part. The two offset grids of raw data are clearly seen in both plots in the overlap regions. The P-V values of the unfiltered data are dominated by spurious data points caused by air bubbles in the glycerol film behind the flat. Form statistics are shown for both plots. The stitched measurement part outline has the correct rectangular form and the correct dimensions, verifying the geometrical aspects of the stitching algorithm. Figure 6.29 shows the stitched and filtered results from the two stitching tests. The raw stitched data were filtered using a 10×10 averaging kernel before computing the filtered statistics.

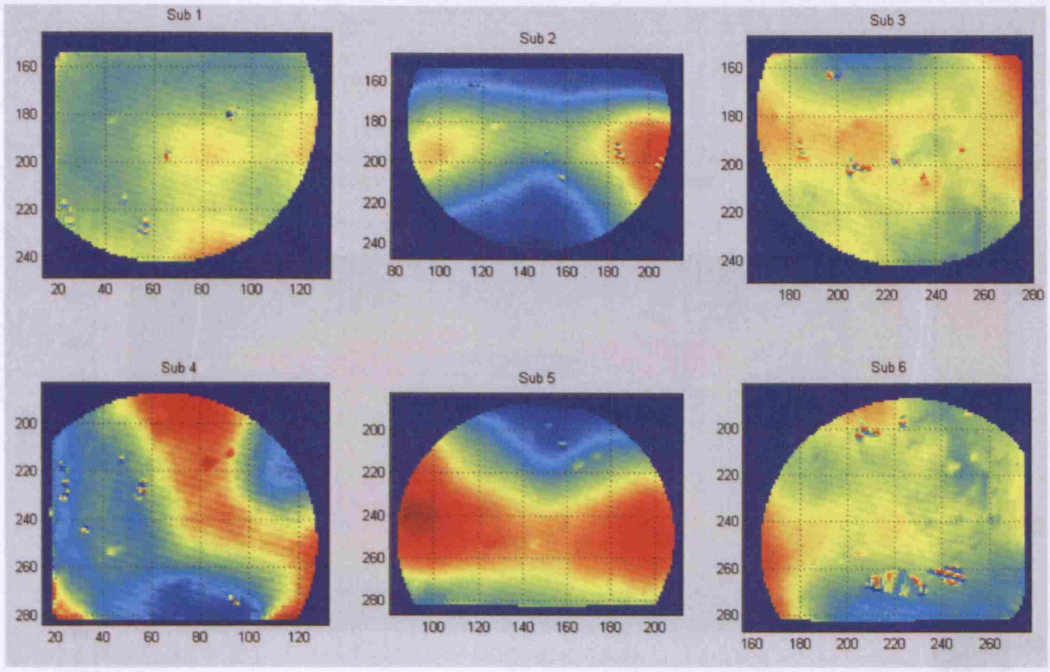


Figure 6.27: Fused-silica sub-apertures from set 1 in their stitched positions.

Figure 6.30 shows the simulated interferograms for the part with a little tip/tilt added for clarity. To obtain a measure of the repeatability, the plots showing the difference between the two measured data sets can be found on figure 6.31.

As no full-aperture test could be made of the part the statistics for the individual sub-apertures and the stitched data are indicated in table 6.5. The data are in 3 sets. Set 1 and set 2 contain the statistics for the individual sub-aperture measurements in columns 1–6, and the statistics for the stitched and filtered data sets in the final data columns. The set labelled 'Difference' is used to benchmark the stitching. Columns 1–6 contain the difference between two measurements of the same sub-aperture, and the final two data columns contain the difference between the two resultant stitched data sets, in both raw-stitched and filtered form.

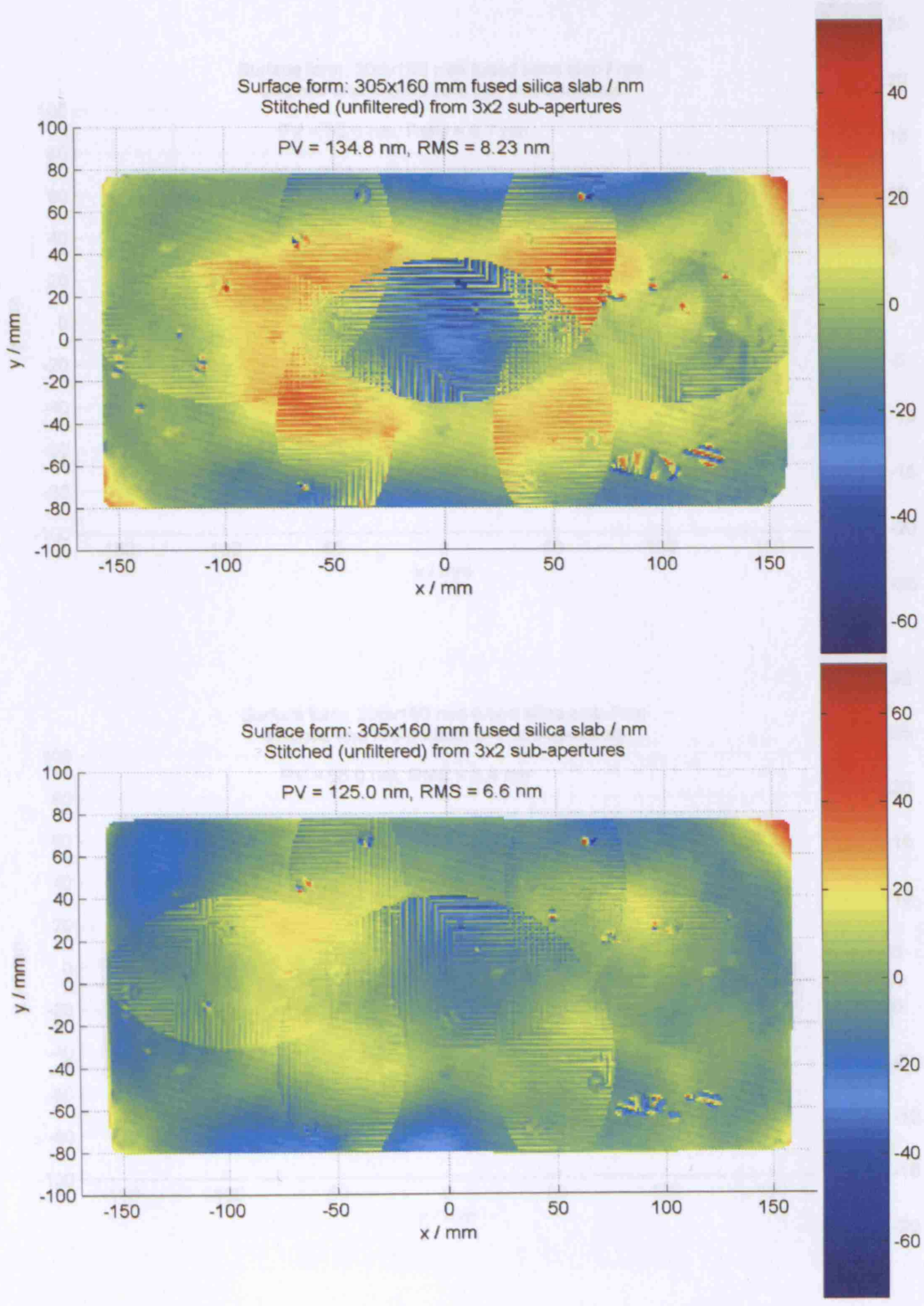


Figure 6.28: Stitched measurement results for the fused-silica flat. Top: set 1. Bottom: set 2.

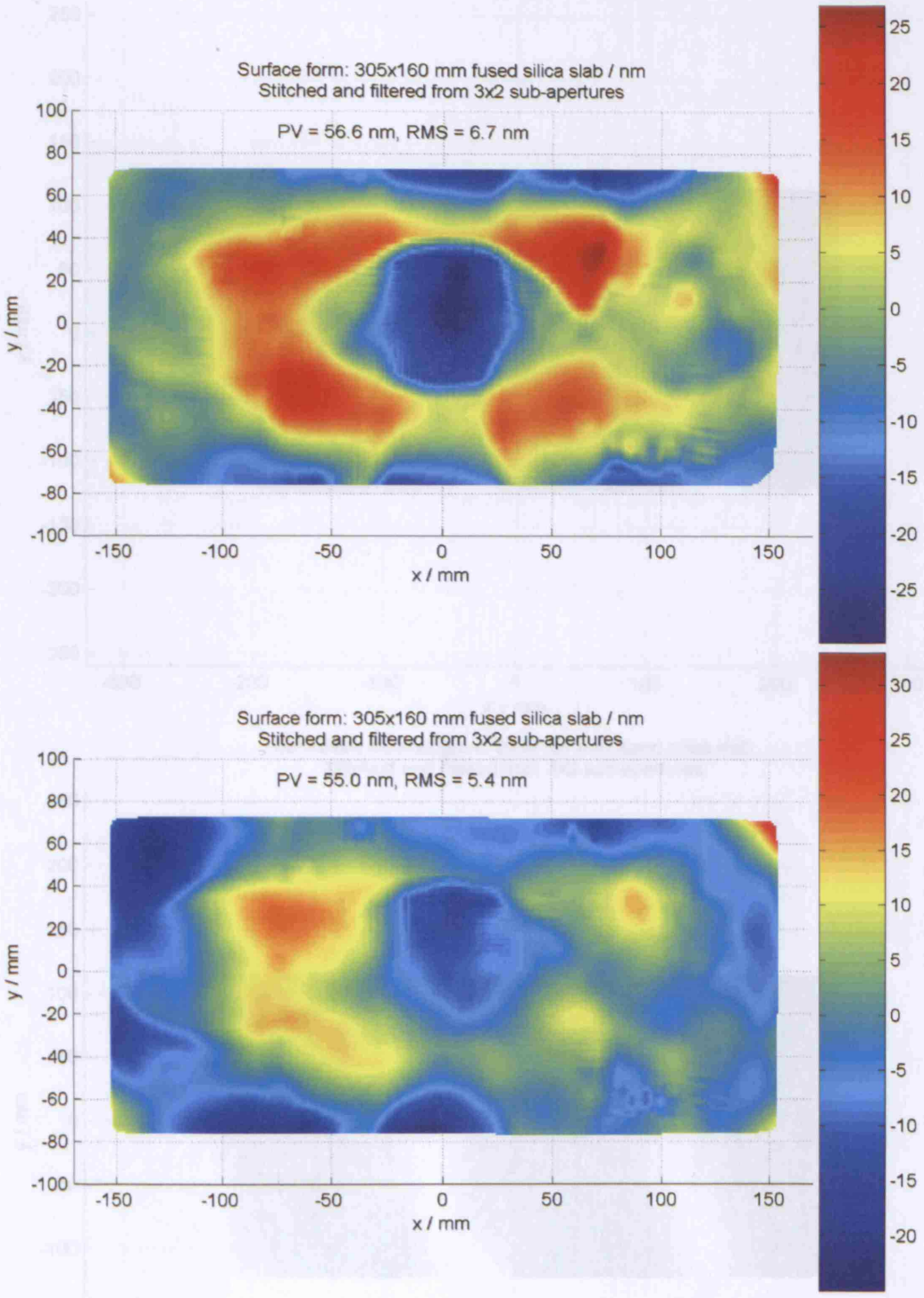
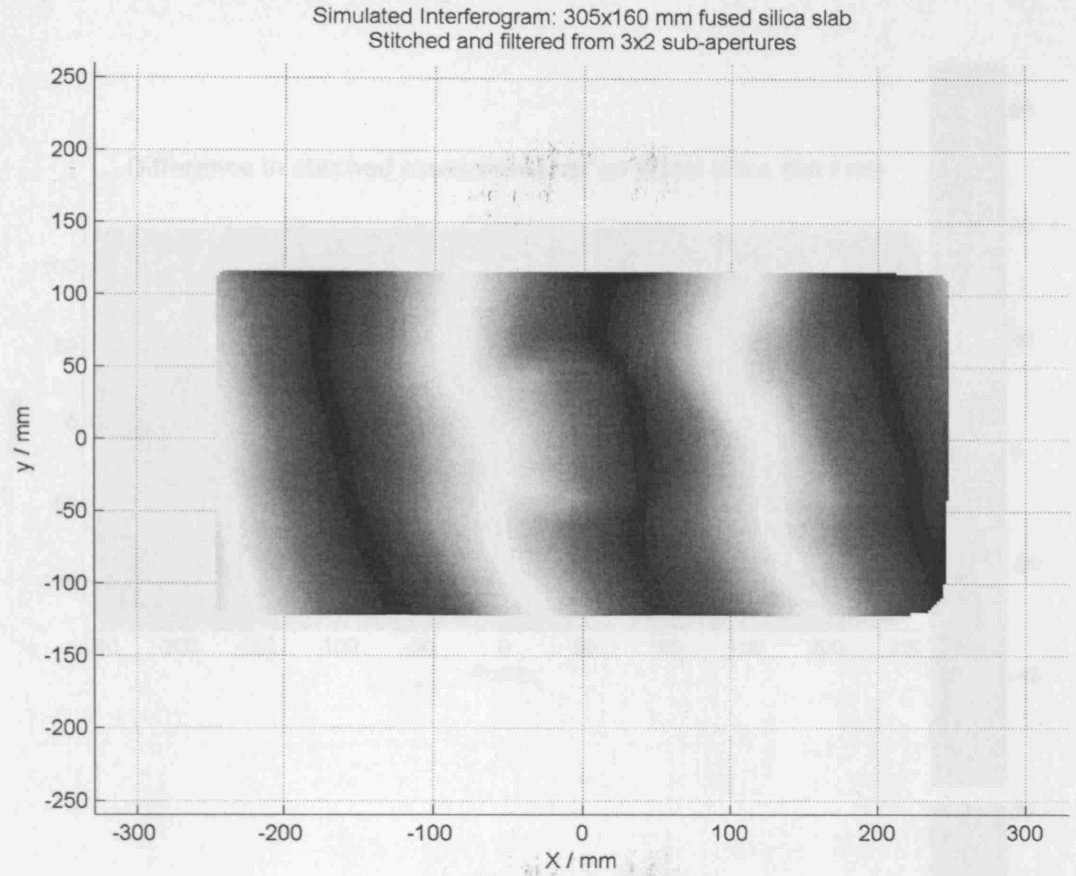


Figure 6.29: Stitched and filtered measurement results for the fused-silica flat. Top: Set 1. Bottom: set 2.



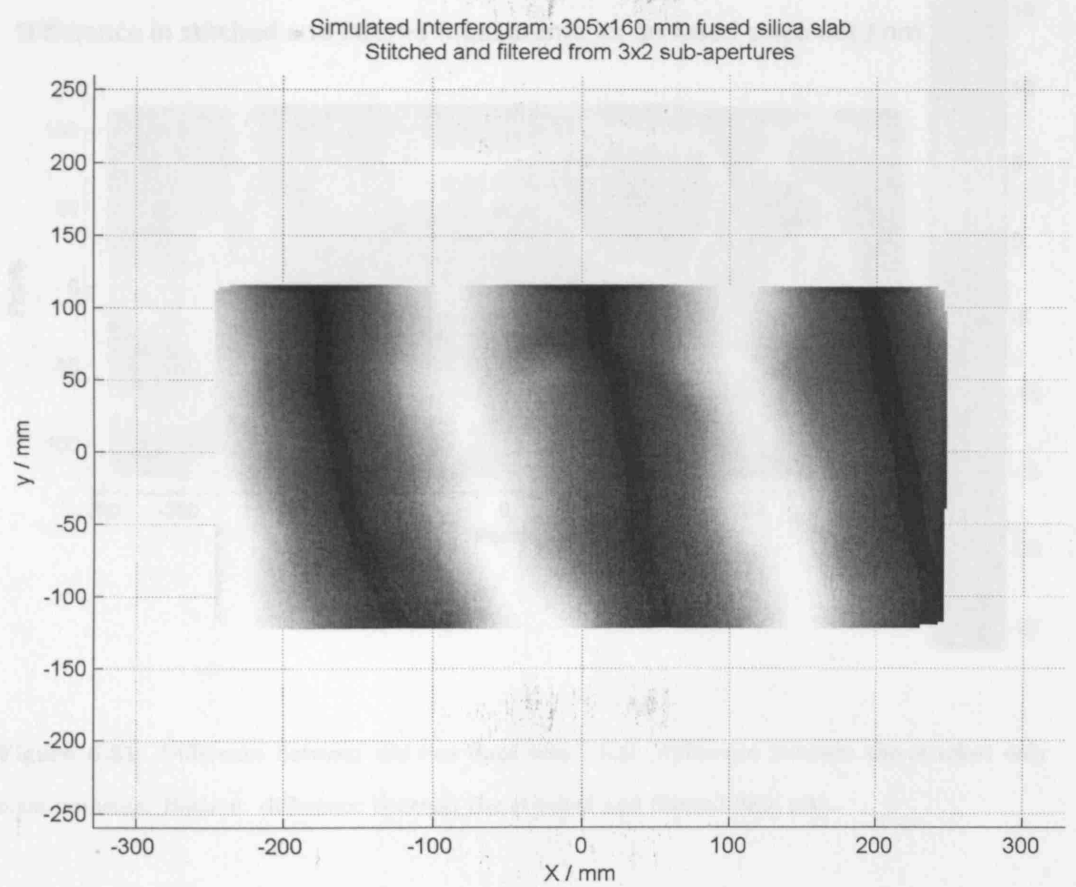


Figure 6.30: Simulated interferograms for the two stitching results on the fused silica flat. Top: set 1. Bottom: set 2.

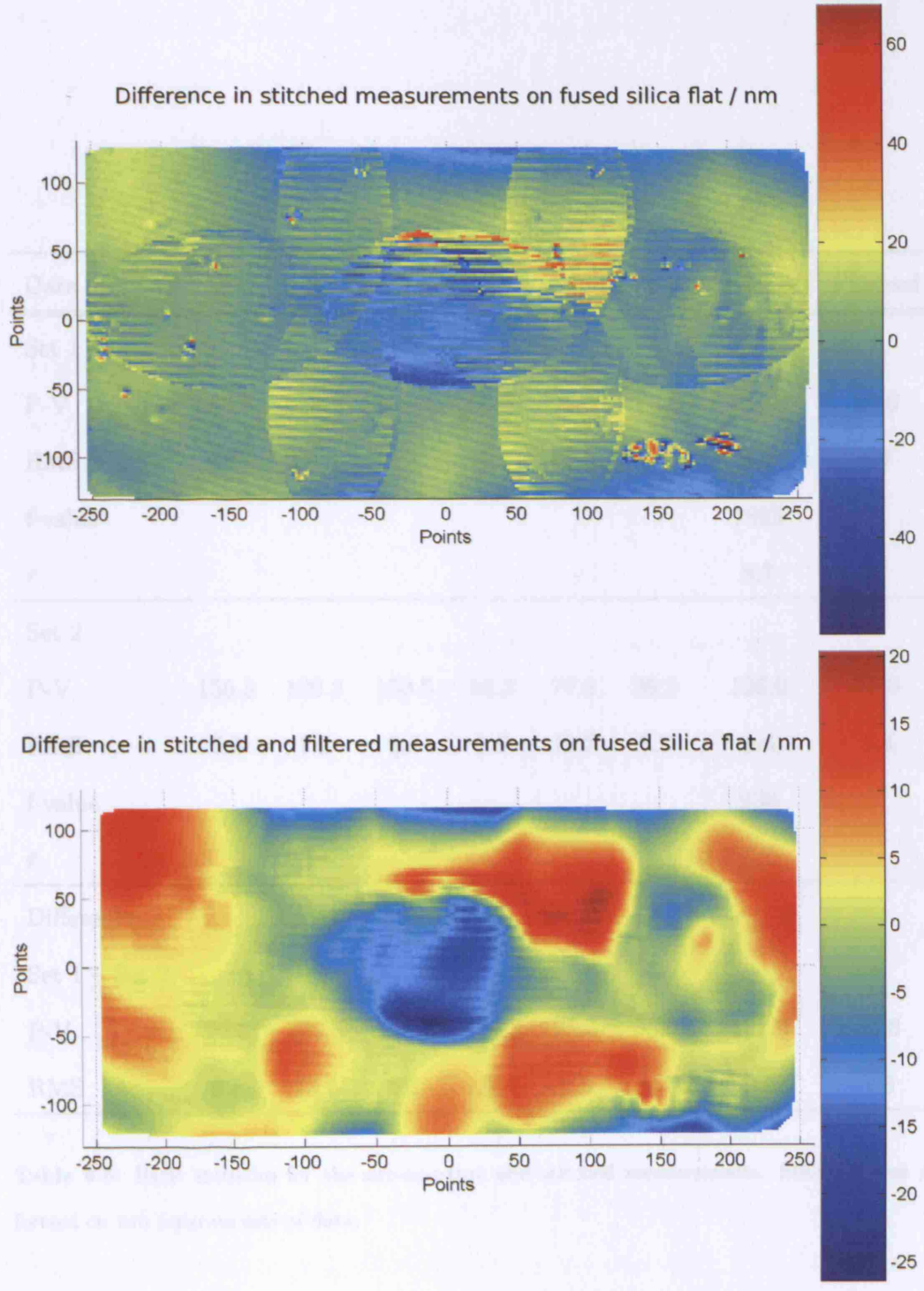


Figure 6.31: Difference between the two data sets. Top: difference between the stitched only measurements. Bottom: difference between the stitched and filtered data sets.

Data	1	2	3	4	5	6	Stitched	Filtered	
Set 1									
P-V	135.8	118.2	165.2	96.1	90.5	103.2	134.8	56.6	nm
RMS	8.6	18.6	4.8	6.9	17.5	7.5	8.2	6.7	nm
f-value							1893		
$\bar{r}$							8.7		
Set 2									
P-V	156.3	109.3	160.5	93.3	77.8	96.2	126.0	55.0	nm
RMS	7.9	7.2	5.6	5.9	12.7	7.7	6.6	5.4	nm
f-value							936		
$\bar{r}$							6.1		
Difference									
Set 1 – Set 2									
P-V	99.3	129.5	163.6	101.3	75.0	63.0	127.6	46.8	nm
RMS	6.8	13.5	5.1	8.2	8.5	4.4	9.9	7.3	nm

Table 6.5: Basic statistics for the sub-aperture and stitched measurements. Stitching was performed on two separate sets of data.

6.5.3 Discussion

The part measured was manufactured to have a surface form error of less than $\lambda/10$ P-V at 632.8 nm. Thus we must expect the P-V form error to be no greater than 63 nm, and expect the RMS form error to be much less than this. The individual sub-aperture measurements have RMS form errors between 4 and 18 nm so it is fair to conclude that the value for the whole part is somewhere close to these values.

The part has a length of 305 mm and no interferometric data exists for the full surface - it is doubted that a 12" aperture interferometer was used to measure the part. It is therefore probable that the part was originally measured against another reference flat using Newton's fringes to establish the P-V form error.

Estimated error

Following from the principles used in section 6.4.3, an estimate of the errors can be made to assess the likely value of $\bar{r}$ from the stitching.

The scaling is 0.61 mm / pixel and we assume an x-y centroid error of 2 pixels since the fiducial sizes on the phasemap were very similar in size to those used for GP_11. Let the total z-error error be 3 nm. This gives an expected residual of 6.9. The object function values at termination were 1893, and 936 for sets 1 and 2 respectively. With 25 fiducial pairs to evaluate, these yield: $\bar{r}$ =8.7 and $\bar{r}$ =6.1 respectively, bracketing the estimated value.

Form

The P-V statistics for the raw measurements and for the unfiltered stitched measurements are often greater than 100 nm. This can be attributed to dirt and scratches on the part, and also to minute bubbles of air in the glycerol film between the fused-silica part and the plate glass. These measurement points inflate the P-V values and are filtered out by the low-pass filtering performed on the data - a standard

procedure in optical interferometry of this type.

A common practice in phase data analysis is to strip pixels from the measurement set. Thus, one, two or three layers of pixels are usually trimmed from around the aperture, and any other 'holes' in the data. The rationale is that edges are usually poorly finished when compared to the bulk of the surface being tested, and that the holes are caused by something on the surface which could result in artificially high (or low) phase values [120]. However, these problems are also alleviated by the filtering process.

An examination of the stitched and filtered data from set 1, in figure 6.29, appears to contain features strongly correlated to the arrangement of the sub-apertures. An examination of figure 6.27 reveals that the features are present in the original sub-aperture data and are not entirely an artefact of the stitching. The data in set 2 stitch with a lower residual, but some small stitching artefacts remain. The overall stitched and filtered statistics are in good agreement with the part specification with the best RMS form error of 5.4 nm, corresponding to approximately $\lambda/100$ at 633 nm.

The error sources to consider are mostly the same as those for GP_11, with the exception of defocus as this is not present here. However, the lower overall RMS values suggest that the remaining errors are the random noise terms from turbulence/convection and vibration, and the systematic errors from the calibration of the transmission flat.

However, the RMS errors between two different measurements of the same part of the surface also have values of this magnitude, as shown in table 6.5. This is suggests that the measurements were made close to limit of the interferometer under the conditions prevailing in the lab at UCL, i.e. the signal-to-error ratio was not great enough to provide a better measurement result.

6.6 Procrustes Algorithm: 100 mm Dia BK-7 flat

6.6.1 Procedure

The $1 \times n$ stitching algorithm was used to stitch 1×2 sub-apertures of data from a 4" diameter BK-7 flat. The flat was polished to $\lambda/20$ RMS and had three fiducials hand drawn on it. The part was measured in two positions one with the part extending beyond the interferometers clear aperture for each. Thus only a fraction of the part was measured. Approximately the centre two-thirds of the part was measured in both tests. Further to this, a full-aperture test was made to enable the stitching to be assessed. The phasemaps were imported into the Matlab software and the sub-apertures were stitched using the Procrustes algorithm.

6.6.2 Results

Figure 6.32 shows the two sub-aperture measurements before stitching, labelled as X and Y. After stitching the data were filtered using the same averaging filter as

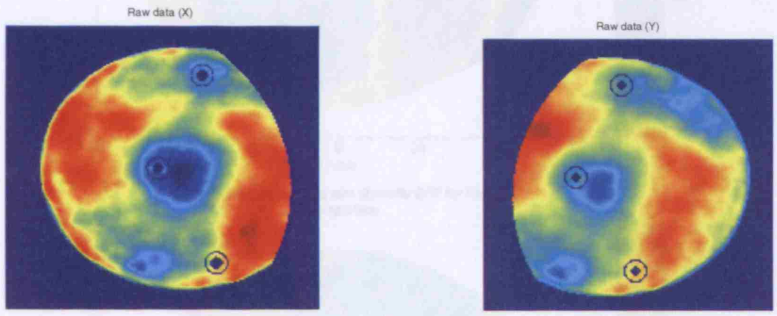


Figure 6.32: Raw sub-aperture measurements for a 4" diameter BK-7 flat ($\lambda/20$ RMS form error). The blue circles are the points used for fiducial centroiding.

for the fused-silica flat. The statistics were also computed for the full-aperture, the stitched and for the stitched and filtered results. Figure 6.33 shows the stitched, stitched and filtered and full measurement results. The peak-to-valley and RMS error measurements were also computed as a further comparison and are shown in

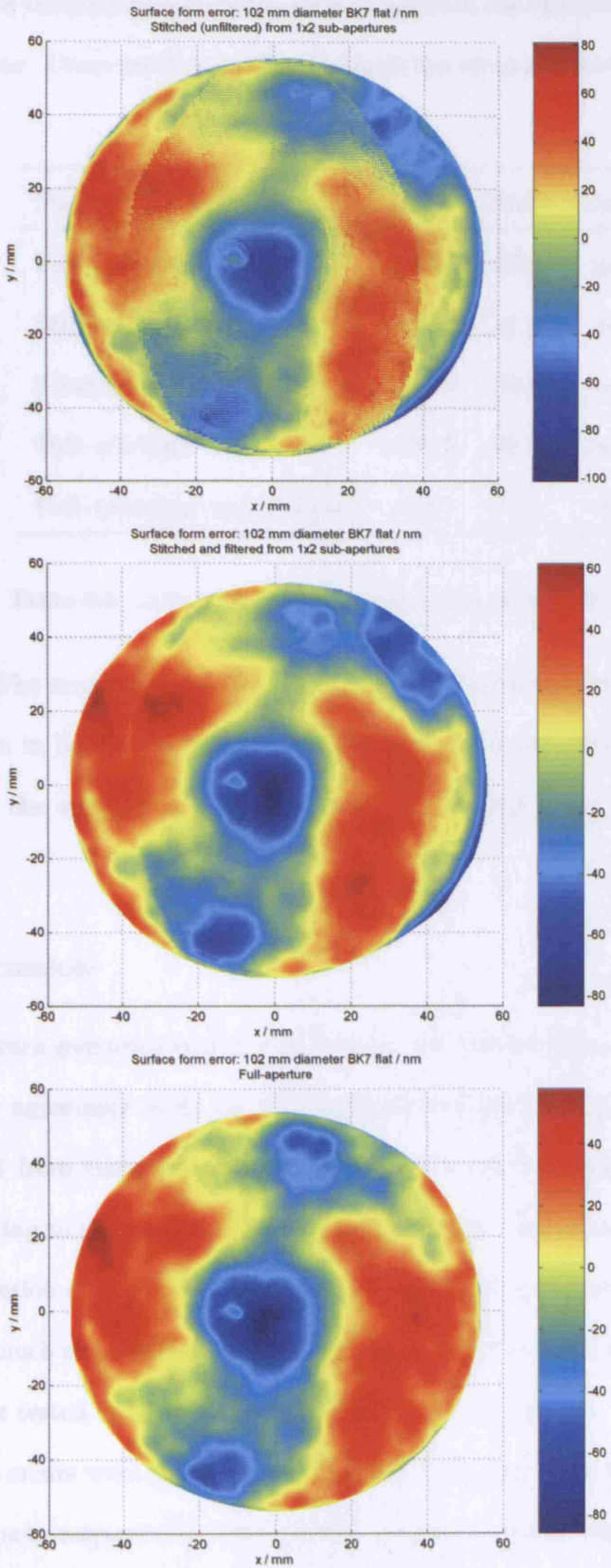


Figure 6.33: Top: stitched measurement result. Middle: stitched and filtered measurement result. Bottom: Full-aperture measurement result.

table 6.6. Note the excellent correspondence between the statistics, in particular for the filtered data. Cross section profiles through the plots are shown in figure 6.34 for

Form Error Statistic	PV	RMS	Units
Full-aperture	143.9	29.5	nm
Stitched	182.4	31.3	nm
Stitched and Filtered	144.9	30.9	nm
Full–stitched	121.2	15.4	nm
Full–(stitched and filtered)	96.3	13.9	nm

Table 6.6: Table of form error measurements for the BK-7 flat.

comparison. The error in the cross sectional data (measured from the full-aperture data) is shown in figure 6.35 As a final comparison the difference between the full-aperture and the stitched data sets was computed and these plots are shown in figure 6.36.

6.6.3 Discussion

The sub-aperture geometry here is very simple and the stitching has been successful giving a close agreement with the full-aperture test giving a similar RMS error to that expected from the measurement errors which are the same as for the fused-silica tests owing to the same mechanical set-up. Again the main factors to improve are the calibration of the interferometer optics and trying to reduce turbulence and vibration as much as possible. For a test on flats, this is easily achieved by placing the part to be tested as close to the interferometer as possible.

The RMS errors were computed as 15.4 and 13.9 nm RMS for the stitched and filtered data sets respectively, attributable mainly to noise and other calibration errors. However, given that the performance of the main optimizing stitcher seems as

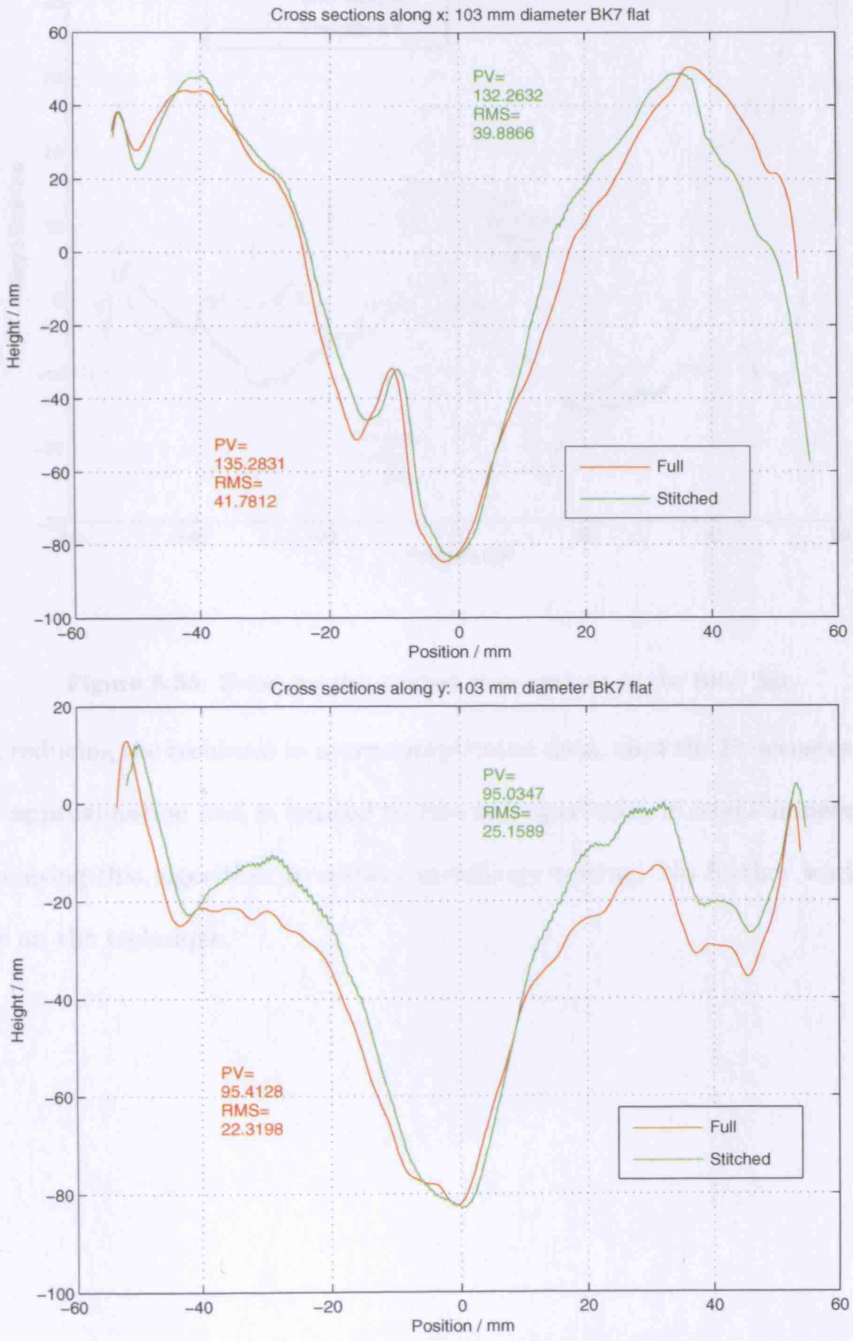


Figure 6.34: Top: cross section through x for BK-7 flat. Bottom: cross section through y for BK-7 flat.

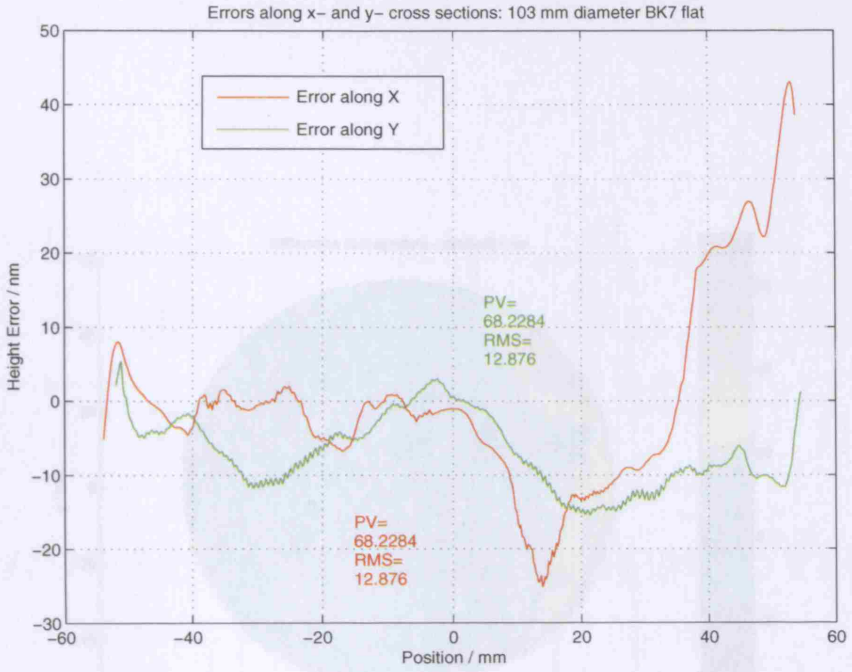


Figure 6.35: Errors for the stitched cross sections on the BK-7 flat.

good at reducing the residuals in more complicated data, that the Procrustes method uses an approximation and is limited to two sub-apertures, it seems unnecessary to consider using this algorithm in optical metrology testing. No further work should be done on the technique.

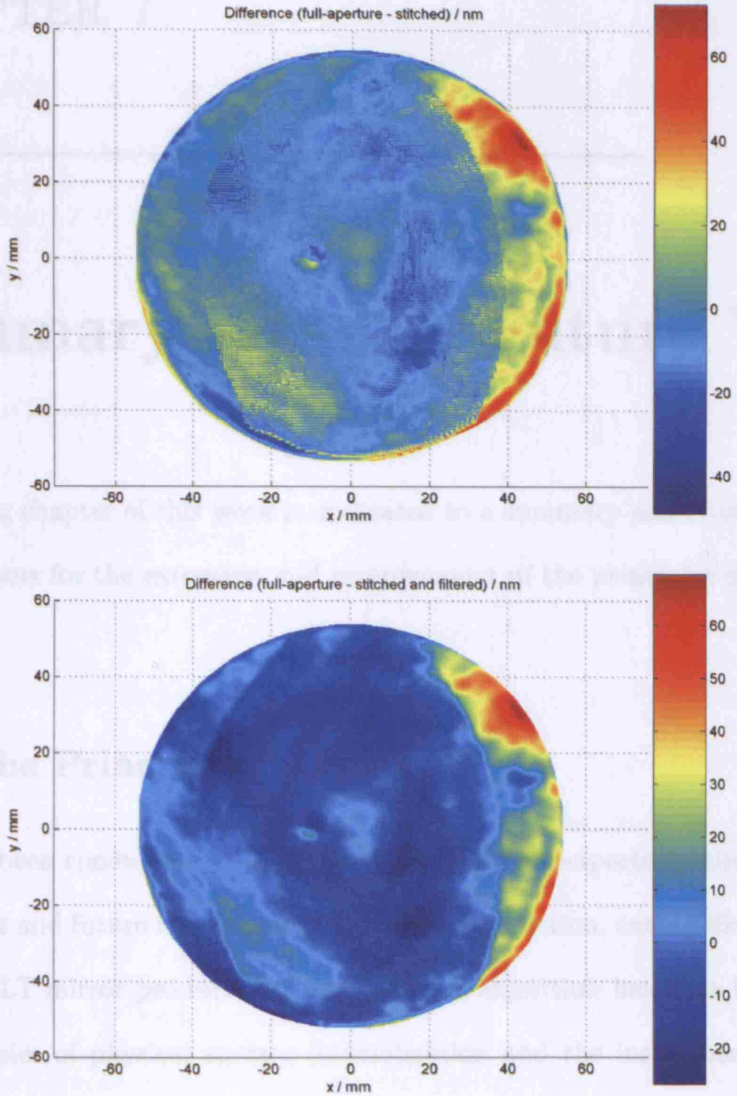


Figure 6.36: BK-7 measurements. Top: difference between the full-aperture and the stitched data set. Bottom: difference between the full-aperture and the stitched and filtered data set.

CHAPTER 7

Summary and the Future Work

The closing chapter of this work is dedicated to a summary and critique of the work and the plans for the extension and improvement of the principles described in this thesis.

7.1 The Principles

Work has been conducted to address the needs for sub-aperture stitching in light of the present and future requirements for optics production, exemplified by the future needs of ELT mirror production. The stitching algorithm has been built up around the principles of physical surface fiducialization and the instruments identified in chapter 2 as a work-around for the non-delivery of the Phase Diversity Wavefront Sensor. The work also avoids stepping on the toes of patented algorithms and is thus eligible for protection via patenting - the author is not aware of any patented algorithms for stitching optical metrology data that use physical surface fiducialization.

The stitching technique has been shown to work, and has stitched data on spheres

and flats with residuals of the same order as the calibration errors of the equipment, and the current limitations of the algorithm.

The processes involved in optical production were reviewed in chapter 1 and these have included both polishing and metrology. In particular, it has been argued that metrology is now the limiting factor in producing many types of optical components, as many modern polishing processes are sufficiently deterministic to produce the required form corrections provided that the metrology can be made accurate enough.

The principle of physical surface fiducialization - the subject of a patent application by NPL - was investigated and a method was found to centroid the fiducials in the metrology data that could be used to compute a set of stitching transformations by a practical approach to numerical minimization.

7.2 Fiducialization

The baseline method of fiducialization involves applying spots of ink to the surface under test using alcohol solvent-based pens in configurations such that the stitching between any two sub-apertures is at least determined by the fiducials alone. This means that at least 3 fiducials are required in any overlap area. This avoids performing least-squares on the whole of the data in the overlap area, and indeed computing which pixels are part of each overlap area. However, it does involve fewer points in the least-squares estimation which will inevitably result in a lower-quality fit.

The presence of a fiducial on the surface means that height data about the surface underneath the fiducial is not available. Centroiding the empty height-data of the fiducials and analysing the surface form around the fiducial has been shown to provide a way of obtaining the x-, y- and z coordinates of a given fiducial.

Much of the concern about this method arises because of the nature of the interferometer data at the fiducial locations and also from the random and systematic errors in the measurement. The x-y centroids of the fiducials are computed on the

positions of approximately 25 data points in a cluster, each with the same numeric value. The centre of mass algorithm therefore weights each pixel equally resulting in a centroid uncertainty of approximately 2 pixels for the analyses in this thesis.

The problems of fiducial alignment are amplified when the resolution of the interferometer's CCD camera is low. In certain circumstances, the resolution unit on the surface under test can approach several mm. Anything much larger is generally problematic if features with spatial frequencies corresponding to the sampling are present on the test surface.

A far better computation of the fiducial's x-y coordinates could be obtained by performing an intensity weighted centre of mass computation on a larger number of pixels.

7.2.1 Revised Fiducialization Method

The Wyko interferometer at UCL is now almost 20 years old and incorporates a low resolution camera with higher noise values compared to modern interferometers. Furthermore, the 4 interferograms captured for the phase-shifting analysis are not available to the user – only one is available as an option. Newer interferometers can make all 4 interferograms available with the output phase array with a well established scale between them. Indeed, the phase values are computed for each CCD pixel so the mapping is generally 1:1 between the 2-D phase arrays.

Modern interferometers typically have cameras with resolutions of 1000×1000 pixels and can be fitted with cameras with up to at least 4000×4000 pixels. The Wyko interferometer used in this work has a useable resolution of approximately 320×320 pixels, so there is much to be gained in terms of resolution by using a more modern system, especially the instantaneous phase-shifting versions.

A modern interferometer was not available for this work, but the following outlines a proposal to improve the fiducial centroiding.

A modern interferometer, such as the Fisba μPhase II, or the 4D Fizcam is used, and the 4 interferograms are averaged on a point by point basis which results in an image of the SUT being formed at the mean level of intensity over the whole of the interferograms. A fully weighted centroiding algorithm can be applied to this image similar to those used in astrometry to provide sub-pixel centroiding. In practice, a threshold would be set and any pixels below this would be considered as fiducial. The algorithm would then identify and centroid the pixels using a centre of mass weighted by the pixel intensities. Algorithms have been shown recently which achieve less than a 10^{th} of a pixel resolution for centroiding stars in CCD images. The centroiding method for the z-coordinate would remain the same since no data would be available for the surface under the ink-spot.

The instantaneous phase-shifting interferometers also have the benefit that the effects of turbulence and stratification are removed from the measurement meaning that the rays in each interferogram all follow the same optical paths, thus reducing the error by an estimated 2-4 nm. Combining the improved centroid with a higher resolution image phasemap of the same surface area will produce a marked improvement of the centroid positions. This modification is therefore recommended for future work.

7.2.2 Automatic Centroiding

This software can be modified to compute the improved centroids, but the recognition procedures remain the same. Upgrading the measurement hardware would allow a higher precision to be obtained as explained in the previous subsection.

7.2.3 Surface Fiducialization

Callender [loc cit.] has explored methods of surface fiducialization using commercial ink-jet printer heads available as OEM parts, but the systems were too expensive to be used in Ph.D. research. Future work with real CNC machines would need

to incorporate sub-aperture and fiducial map design software with CNC software. Completion of this software would represent one step towards a fully automated system.

7.3 Geometric Transformations

Chapter 3 showed how to construct a set of general transformation matrices to perform the sub-aperture stitching after measurement. The principles of homogeneous coordinates, affine transformations and compound matrices were used and a generic transformation matrix **T** was developed whose elements are found by a least-squares numerical optimization.

The transformations were applied to only the fiducial coordinates in the sub-apertures to minimize the number of points to be transformed, and thus the CPU time used, on each object function evaluation in the optimizer's iterative routine. A device called the fiducial matrix was developed that stored the locations of the fiducials in the sub-apertures and also allowed the number and configuration of the sub-apertures to be automatically found.

It was also shown how the fiducial matrix could be used to compute the number of fiducial residuals to sum, and hence, how final object function values could be estimated by considering the errors in the measurement and centroiding. The reverse also applies such that a mean fiducial residual can be evaluated for each stitching run.

7.4 Optimization and Data Output.

The tasks to be completed before the optimization process were considered in chapter 4, then the optimization itself. Further optimizations were shown, one as a drop-in replacement for the main program, and one for the special case of stitching n×1 sub-apertures in a sequential fashion whose further use is not advocated.

The theme of the optimization was concerned with using a commercially available optimization routine rather than attempting to embody most of the current knowledge of numerical minimization into a custom algorithm. Matlab's optimization toolbox was used because it is very comprehensive and also because Matlab is the environment used for other software by the industrial sponsors, Zeeko. The use of Matlab means that the software can be incorporated into the Zeeko software suite without having to resort to re-coding or purchasing an alternative environment. The main work involved the use of an optimization using a quasi-Newton line search and a Hessian updating system and forms the main stitching algorithm which allows any geometrical arrangement of overlapping sub-apertures to be stitched. A simplex algorithm was also described that could directly replace the main one and was used in a low-dimensional stitch of simulated data to provide a stitching solution to a two sub-aperture problem with fiducial residuals of the order of 10^{-26} units.

7.5 The Swing-Arm Profilometer

The NPL/UCL Swing-Arm Profilometer (SAP) was discussed and how it could be used as a platform for stitching metrology data. The chapter contained some tentative ideas for future work involving the possible mounting of an interferometer on the swing-arm. It is hoped that the SAP will be used for future experiments in stitching with this work as it provides a baseline CNC technology system for scanning an interferometer over a spherical surface and maintaining that device perpendicular to the surface at all times.

7.6 Stitching Tests

The accurate centroiding of fiducials is central to successful stitching and reducing the errors is a necessary objective of future work. The stitching results presented in the previous chapter have been shown to be successful in the limit of these mea-

surement and fiducial centroiding errors, and proposed methods for their reduction has been discussed in the previous sections.

The results also show the need for longer optimization times as the number of sub-apertures increases, as expected. This does not really pose a problem where a small number of sub-apertures is to be encountered, but when we think of ELT segments, we may be faced with over 100 sub-apertures per segment or more - >600 variables to optimize, clearly a CPU intensive task. The advantage of this work by operating only on fiducials during the optimization process should now be clearer. Transforming 100 sub-apertures 600 times per iteration[1] would be very time consuming. Operating only on fiducials is computationally efficient in comparison, in terms of time and memory use.

So where will the real improvements be made in future work? It seems likely that these will occur in decreasing the measurement errors and making the optimizer more efficient.

7.6.1 Decreasing the Errors

High quality calibration of the transmission elements was not available for this project because of time and cost. For future work the sphere calibration of Parks, described in chapter 6 should be used to remove the interferometer systematic errors. For example, Parks in [121] showed a calibration of a good quality f/3.3 transmission sphere with an RMS form error of 0.017 waves at 633 nm (10.8 nm). Clearly, errors of this magnitude over the field of view of the interferometer will affect the stitching accuracy if they are not mapped. In this work, the transmission spheres and flats were calibrated with reference objects whose form was guaranteed only to $\lambda/10$ or $\lambda/20$ P-V, whose absolute surface form was not known to the order of a few nm. Clearly, the form errors of the transmission elements will be present in the stitched

[1] Remembering that the finite differencing perturbs each design variable in turn to estimate the gradient of the object function value, and that there are 6 variables per sub-aperture

data.

The repeatability for this work has been shown to be approximately 3.7 nm by consideration of a series of repeated measurements of the same part over path lengths less than 1 m. This number was established in central London in a laboratory just 60 m from passing underground trains with a data acquisition time of 167 ms. Transmission spheres therefore need to be well characterised to leave room just for the random component of the errors and the procedure of Parks appears to provide the accuracy required.

It is therefore expected that in a stable environment that errors from vibration can be reduced to the level of 1 nm. Good calibration of transmission spheres should be achieved such that the RMS error of the transmission sphere calibration is less than the repeatability of the measurement as suggested by Griesmann [122]. This is generally achieved by taking sufficient measurements on the calibration sphere in Parks' method.

Errors from defocus terms were discussed in the results section. Indeed, it seems that there is work to be done here on correcting defocus terms in non-circular apertures. This is suggested as yet another possible improvement to the stitching algorithm, but it can be envisaged that computing a defocus term for each sub-aperture several times per iteration will also be very time consuming.

7.6.2 Speeding Up the Optimization

The main methods for speeding up the optimizer are as follows:

1. Reducing the data passing to the optimizer at each iteration.

2. Providing a numerical routine for evaluation of the object function gradient.

3. Buy a faster computer.

In (1) this can be achieved by avoiding passing the sub-aperture data to the object function at each iteration. This could involve making a copy of just the fiducial co-ordinates in each sub-aperture and passing only these and would require the fiducial matrix to be recomputed to account for the change.

One of the best ways to reduce an optimization time of this kind is to provide an analytical gradient vector. Providing analytical gradients in (2) above is more difficult than step (1) above, but would have a larger effect on the optimization time. This is evident from the object function routine in appendix G. An investigation into the provision of such gradients is another suggestion for further work, but would require a complete re-expression of the object function. Of course, one could just buy a faster computer...

7.7 Measuring Aspheric Surfaces

The measurement of aspheric surfaces presents several extra problems in stitching metrology data and the techniques developed in this thesis will need to be augmented to cope with the additional parameters of the aspheric surface. This section is devoted to exploring some of the points to be addressed when the part to be measured is an ELT segment.

Some of the extra points to consider are:

1. The radius of curvature changes over the part which must be accounted for in making the measurement.

2. The aspheric departure over any sub-aperture measurement must not be too great for the interferometer in use.

3. Under the same measurement conditions as a spherical part, the error budget is the same except that the re-trace errors become larger in the absence of a null corrector.

We can begin by considering the context of measuring an ELT segment from the ESO telescope. We first consider a spherical form on the segment with a radius of curvature of 84 m. The most practical way performing sub-aperture stitching on this segment is by equipping the measurement interferometer with a diverging transmission element. The confocal point of the system is virtual (behind the interferometer) and the distance to the surface under test can be minimized by allowing the radius of curvature of the test surface to closely match that of the part to within a few 10s of mm. The interferometer can then be scanned over the part by a CNC machine system such as the NPL/UCL swing-arm profilometer. Figure 7.1 illustrates a full-aperture test at the centre of curvature versus a sub-aperture test with a virtual confocal point.

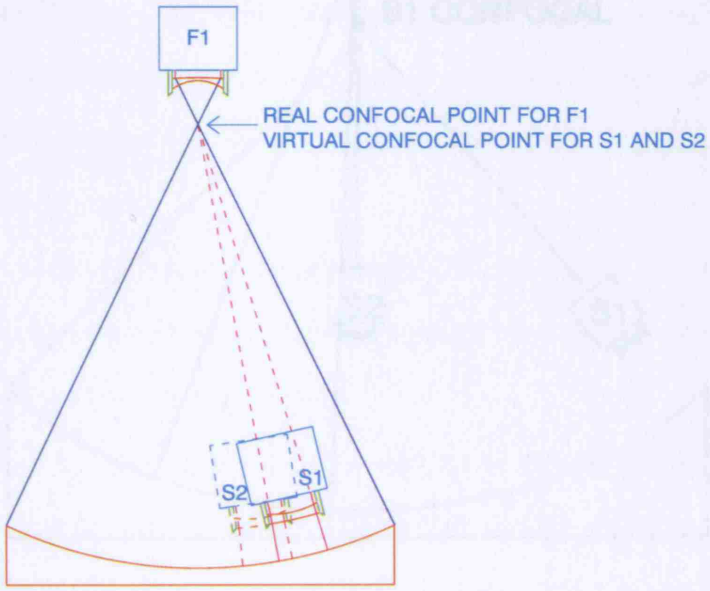


Figure 7.1: Testing an ELT segment. F1 is an interferometer performing a full-aperture test of the segment using a converging transmission element. S1 and S2 represent interferometers performing sub-aperture tests on the same segment using weakly diverging transmission elements. The long path length can be avoided using S1 and S2 by designing a Fizeau transmission element whose reference surface has a radius of curvature close to that of the part under test.

If we now consider aspherizing the segment to the ESO specifications, we will

have an off-axis asphere in the form of a prolate ellipsoid. We will now have to servo the interferometer to track the local centre of curvature across the part. If we do not do this, two things will occur.

1. The fringe density may become too high for the interferometer to cope with.

2. The retrace errors will become large as a consequence of (1).

Tracking the radius changes allows the sub-aperture measurement to be made as close as possible to null thus reducing the retrace errors caused by the asphericity. The process is illustrated in figure 7.2. The interferometer requires tip, tilt

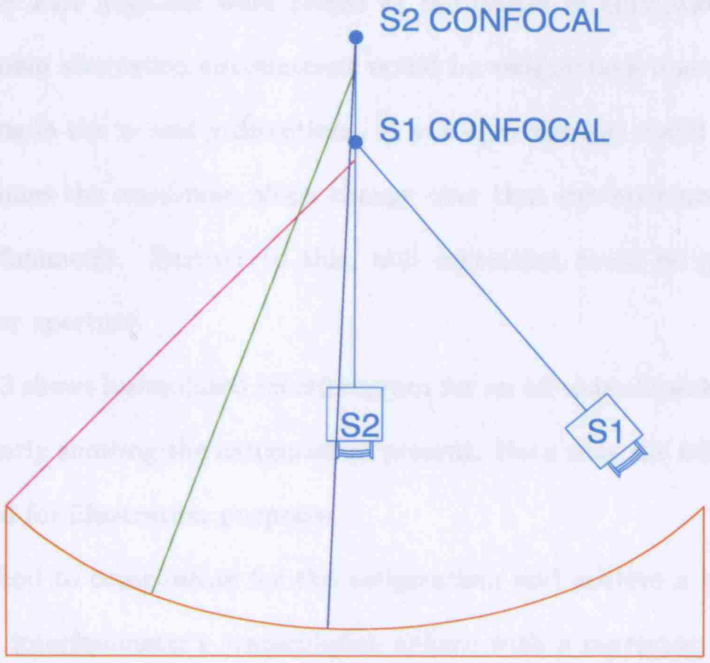


Figure 7.2: A proposal for testing an ellipsoidal ELT segment: exaggerated schematic. S1 is measuring the part over a region where the radius is smaller than that where S2 is measuring. The interferometer therefore needs to be servoed to achieve the best set of fringes in each case. The green, blue and magenta lines show surface normals whose upper endpoints would be where the interferometer-part confocal point would be located.

and linear translation adjustment along its optical axis to achieve the best set of fringes. The tip/tilt and translation axes can be equipped with precision encoders

or interferometers to monitor the changes which could be used to perform initial modifications to the phasemaps before stitching.

Much of the stitching formalism developed in this work remains unchanged:

1. Design an appropriate fiducialization scheme for the surface to be measured.
2. Fiducialize the surface.
3. Position the interferometer to measure each sub-aperture.
4. Stitch the measurements with an improved stitching algorithm.

If such an ELT segment were tested at the centre of curvature of its best-fit sphere, the main aberration encountered would be astigmatism due to the nature of the curvatures in the x- and y-directions. A sub-aperture size could be chosen such that at all times the maximum slope change over that sub-aperture is measurable by the interferometer. Further to this, null correction could be provided at the interferometer aperture.

Figure 7.3 shows a simulated interferogram for an off axis ellipsoid from its best-fit sphere clearly showing the astigmatism present. Note that the fringe spacing has been adjusted for illustration purposes.

One method to compensate for the astigmatism and achieve a better null is to augment the interferometer's transmission sphere with a corrector. A simple idea is to equip the interferometer with two extra lenses, each containing an astigmatic term of opposite sign. The lenses could be designed to rotate relative to each other and to rotate as a whole relative to the interferometer. Therefore a given quantity of astigmatism could be programmed into the lenses by rotating them relative to each other. Furthermore, the whole assembly could be rotated relative to the interferometer to align the correction to the local surface under test. Such an idea has already been posed as a compensation method for a full aperture test of such ELT segments [73].

Figure 7.3: Simulated interferogram of an ESO ELT segment. The fringe scaling has been increased otherwise the plot would contain too many fringes to display. The interferogram has been generated by comparing the ELT segment's surface to its best-fit sphere. The dominant aberration is astigmatism, although other modes are present.

After the sub-aperture measurements have been taken the stitching can proceed but with some alterations to the stitching algorithm due to the following points:

1. The surface area imaged in each sub-aperture will vary slightly due to the change in radius of curvature.

2. A sub-aperture will have to be set as the reference sub-aperture. This will have a fixed tip/tilt and a fixed scale in the three Cartesian coordinate directions.

Item (1) can be accommodated in the stitching software by the introduction of extra parameters to be optimized to include small-magnitude scaling along the measurement's principal axes. This would allow for a group of fiducials to be matched more closely between sub-apertures. In theory more complicated mappings could be introduced that more accurately represent the design surface form. As a simple example, equation 3.18 could be modified to allow the scaling parameters as **TS**, where **S** is an homogeneous $\mathbb{R}^3$ matrix with scaling in x and y, to give

$$\mathbf{TS} = \begin{pmatrix} S_x \cos\psi & -S_x \sin\psi & 0 & S_x(t_x \cos\psi - t_y \sin\psi) \\ S_y \sin\psi & S_y \cos\psi & 0 & S_y(t_x \sin\psi + t_y \cos\psi) \\ A & B & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.1}$$

Item (2) is somewhat simpler to implement. We could choose a central sub-aperture to be the reference and stitch everything onto this. This sub-aperture could be scaled with reference to equation 4.2 to provide the correct scaling of the phasemap.

A further improvement in the stitching would be to monitor the absolute distance to the surface under test with a distance measuring interferometer. This distance can be used to compute the local radius of curvature of the part which can be fed into the optimizer to tie down the scaling parameters.

With these, it becomes possible to contemplate the measurement of aspheric parts and a system incorporating these features would of course require a CNC

system to provide large-scale servoing across the part - a Zeeko IRP machine or the swing-arm profilometer for example.

7.8 Final Remarks

This work has been about finding a new approach to stitching metrology data and has combined some novel approaches and worked around equipment non-availabilities and peculiarities to provide what is hoped to be an extensible system that works within the measurement errors present in the metrology.

It provides a software basis for further development into an on-machine metrology system utilizing sub-aperture stitching and CNC technology. Of course, it will not be a stand-alone algorithm. It will be part of a large process involving many calibrations, procedures and systems, some of which are covered in this thesis, and which will hopefully be used after further development and improvement to form a toolkit of parts that can be bolted together to solve a particular metrology problem.

It is envisaged that the work will be extended in the very near future, with particular regard to reducing errors and the asphere problem.

REFERENCES

[1] M. Otsubo, K. Okada, and J. Tsujiuchi. Measurement of large plane surface shapes by connecting small-aperture interferograms. *Optical Engineering*, 33(2):608–613, February 1994.

[2] Wikipedia. History of Optics. http://en.wikipedia.org/wiki/History_of_optics, December 2006.

[3] Encyclopaedia Britannica Online Articles. Ptolemy. http://search.eb.com /eb/article-9061778, March 2007.

[4] Encyclopaedia Britannica Online Articles. Almagest. http://search.eb.com /eb/article-9005845, March 2007.

[5] Encyclopaedia Britannica Online Articles. Kindi, Ya'qub Ibn Ishaq as-Sabah, al-. http://search.eb.com/eb/article-9045485, March 2007.

[6] Wikipedia. Al-kindi — Wikipedia, The Free Encyclopaedia, 2007. [Online; accessed 6-March-2007].

[7] Wikipedia. Ibn Sahl — Wikipedia, The Free Encyclopaedia, 2007. [Online; accessed 6-March-2007].

[8] R. Rashed. A pioneer in anaclastics: Ibn Sahl on burning mirrors and lenses. *Isis*, 81:464–491, 1990.

[9] Encyclopaedia Britannica Online Articles. Ibn al-Haytham (Alhazen). http://search.eb.com/eb/article-217919, March 2007.

[10] Encyclopaedia Britannica Online Articles. Bacon, Roger. http://search.eb.com/eb/article-536, March 2007.

[11] Encyclopaedia Britannica Online Articles. Copernicus, Nicolaus. http://search.eb.com/eb/article-9105759, March 2007.

[12] Encyclopaedia Britannica Online Articles. Lippershey, Hans. http://search.eb.com/eb/article-9048449, March 2007.

[13] Justus Sustermans. Galileo Galilei. National Maritime Museum, Greenwich http://www.nmm.ac.uk/mag/pages/mnuExplore /PaintingDetail.cfm?ID=BHC2700, 1639.

[14] Encyclopaedia Britannica Online Articles. Newton, Sir Isaac. http://search.eb.com/eb/article-12245, March 2007.

[15] Wikipedia. Snell's law — Wikipedia, The Free Encyclopaedia, 2007. [Online; accessed 6-March-2007].

[16] Lasers and Optics. A brief history of optics on the www. http://members.aol.com/WSRNet/D1/hist.htm, March 2007.

[17] Encyclopaedia Britannica Online Articles. Optics. http://search.eb.com, March 2007.

[18] S. Hampton. Glassmaking in antiquity. University of North Carolina: http://www.unc.edu/courses/romtech/public/content/arts_and_crafts /Susan_Hampton/Roman_Glass.html, Retrieved 2006.

[19] Encyclopaedia Britannica Online Articles. Silvering. http://search.eb.com /eb/article-9067821, March 2007.

[20] Encycloaedia Britannica Online Articles. Optics. http://search.eb.com /eb/article-37962, March 2007.

[21] D. Malacara *et al.* *Optical Shop Testing.* Wiley, 1st edition, 1978.

[22] Precision-Optical Engineering. Definition of an aspheric surface. Technical report, P-O E, 2005.

[23] D. D. Walker. Specifications for metrology for various surfaces. Technical report, University College London, April 2004. Personal communication.

[24] L. Piegl and W. Tiller. *The NURBS Book.* Springer-Verlag, 2nd edition, 1997.

[25] M. Rowan-Robinson. Astronomy in the UK: A review of UK astronomy and space science. Royal Astronomical Society, 2007.

[26] H. Benson. *University Physics.* Wiley, 1995.

[27] GMT Group. GMT Conceptual Design Report. Technical report, GMT, 2007.

[28] G. Angeli and S. Roberts. Thirty meter telescope: Observatory Architecture Document. Technical report, TMT, 2007. Document: TMT.SEN.DRD.05.002.CCR16.

[29] I. M. Hook, editor. *The Science Case for the European Extremely Large Telescope.* Holly Benson Communications, 2005.

[30] D. D. Walker *et al.* Design study report: The primary and secondary mirrors for the proposed Euro50 telescope. Technical report, Euro50 Project, April 2002.

[31] E. Brunetto, P. Dierickx, R. Gilmozzi, M. Le Louarn, F. Koch, L. Noethe, Ch. Verinaud, and N. Yaitskova. Progress of ESO's 100 m OWL optical telescope design. In A. L. Ardeberg and T. Andersen, editors, *Second Bäckaskog Workshop on Extremely Large Telescopes*, volume 5382, pages 159–168. SPIE, September 2003.

[32] ESO. The Messenger, March 2007.

[33] David Walker. Personal communication, 2007.

[34] Personal communication with UPS group, Optic Technium N. Wales., 2007.

[35] J. Lubliner and J. E. Nelson. Stressed mirror polishing. 1: A technique for producing nonaxisymmetric mirrors. *Applied Optics*, 19(14):2332–2340, July 1980.

[36] J. E. Nelson, G. Gabor, L. K. Hunt, J. Lubliner, and T. S. Mast. Stressed mirror polishing. 2: Fabrication of an off-axis section of a paraboloid. *Applied Optics*, 19(14):2341–2352, July 1980.

[37] J. Zimmerman. Computer controlled optical surfacing for off-axis aspheric mirrors. In *Advanced Technology Optical Telescopes*, volume 1236, pages 663–668. SPIE, 1990.

[38] H. M. Martin, D. S. Anderson, J. R. P. Angel, R. H. Nagel, S. C. West, and R. S. Young. Progress in the stressed-lap polishing of a 1.8 m f/1 mirror. In *Advanced Technology Optical Telescopes*, volume 1236, pages 682–690. SPIE, 1990.

[39] Y. Verma, A. K. Chang, J. W. Berrett, K. Futtere, G. J. Gardopee, J. Kelley, T. Kyler, J. Lee, N. Lyford, D. Proscia, and P. R. Sommer. Rapid damage-free shaping of silicon carbide using Reactive Atom Plasma (RAP) processing. In

E. Atad-Ettedgui, J. Antebi, and D. Lemke, editors, *Optomechanical Technologies for Astronomy*, volume 6273. SPIE, 2006.

[40] C. Fanara, P. Shore, J. Nicholls, N. Lydford, P. Sommer, and P. Fiske. A new reactive atom plasma technology (RAPT) for precision machining: the etching of ULE optical surfaces. In E. Atad-Ettedgui, J. Antebi, and D. Lemke, editors, *Optomechanical Technologies for Astronomy*, volume 6273, pages 62730B–1–8. SPIE, May 2006.

[41] L. N. Allen and H. W. Romig. Demonstration of an ion figuring process. In *Advanced Optical Manufacturing and Testing*, volume 1333. SPIE, 1990.

[42] Richard G. Bingham, David D. Walker, Do-Hyung Kim, David Brooks, Richard Freeman, and David Riley. A novel automated process for aspheric surfaces. In Robert E. Fischer, R. Barry Johnson, Warren J. Smith, and William H. Swantner, editors, *Current Developments in Lens Design and Optical Systems Engineering*, volume 4093, pages 445–450. SPIE, 2000.

[43] David D. Walker, Richard Freeman, Gerry McCavana, Roger Morton, David Riley, John Simms, David Brooks, Eugene Kim, and Andrew King. Zeeko/UCL process for polishing large lenses and prisms. In Richard G. Bingham and David D. Walker, editors, *Large Lenses and Prisms*, volume 4411, pages 106–111. SPIE, 2002.

[44] D. Walker, A. Beaucamp, C. Dunn, R. Freeman, A Marek, G. McCavana, R. Morton, and D. Riley. First results on free-form polishing using the Precessions process. In *Winter Topical Meeting on 'Free-Form Optics: Design, fabrication, Metrology, Assembly'*. ASPE, February 2004.

[45] H. S. Yang. *Developments in Stylus Profilometry.* PhD thesis, University College London, 2000.

[46] Yang Ho Soon. *Developments in Stylus Profilometry.* PhD thesis, University College London, 2000.

[47] M. V. R. K. Murty. *Optical Shop Testing*, chapter 1, pages 19–36. Wiley, 1st edition, 1978.

[48] J. C. Wyant and V. P. Bennett. Using computer generated holograms to test aspheric wavefronts. *Applied Optics*, 11(12):2833–2839, December 1972.

[49] J. Primot. Theoretical description of a Shack-Hartmann wavefront sensor. *Optics Communication*, (222):81–92, 2003.

[50] F. Roddier, C. Roddier, and N. Roddier. Curvature sensing: a new wavefront sensing method. In *Advanced Technology Optical Telescopes*, pages 203–209. SPIE, 1988.

[51] C. Roddier, F. Roddier, A. Stockton, and A. Pickles. Testing of telescope optics: a new approach. In *Advanced Technology Optical Telescopes IV*, volume 1236, pages 756–766. SPIE, 1990.

[52] P. M. Blanchard and A. H. Greenaway. Multi-plane imaging with a distorted diffraction grating. *Tech Report*, 1999.

[53] P. M. Blanchard and A. H. Greenaway. Simultaneous multi-plane imaging with a distorted diffraction grating. *Applied Optics*, 38(32):6692–6699, November 1999.

[54] P. M. Blanchard, D. J. Fisher, S. C. Woods, and A. H. Greenaway. Phase-diversity wave-front sensing with a distorted diffraction grating. *Applied Optics*, 39(35):6649–6655, December 2000.

[55] S. C. Woods and A. H. Greenaway. Wave-front sensing by use of a Green's function solution to the intensity transport equation. *J. Opt. Soc. Am. A*, 20(3):508–512, March 2003.

[56] H. I. Campbell, S. Zhang, and A. H. Greenaway. Generalized phase diversity for wave-front sensing. *Optics Letters*, 29(23):2707–2709, December 2004.

[57] S. Arnold. Interferometric null testing of aspheres. SPIE OE Magazine, August 2002.

[58] T. Andersen and A. Ardeberg *et al.* Euro 50. Technical report, Lund Observatory, August 2003. Design study of a 50m adaptive optics telescope.

[59] A. Efstathiou. *Design Considerations for a Hybrid Swing-Arm Profilometer to Measure Large Aspheric Optics.* PhD thesis, University College London, 2007. In progress.

[60] J. C. Gower and G. B. Dijksterhuis. *Procrustes Problems.* OUP, 1st edition, 2004.

[61] S. C. Jensen, W. W. Chow, and G. N. Lawrence. Subaperture testing approaches: a comparison. *Applied Optics*, 23(5):740–745, March 1984.

[62] W. Chow and G. Lawrence. Method for subaperture testing interferogram reduction. *Optics Letters*, 8(9):468–470, September 1983.

[63] J. Thunen and O. Kwon. Full aperture testing with sub-aperture test optics. In *Proceedings of SPIE: Wavefront Sensing*, volume 351, pages 19–27. SPIE, February 1983.

[64] M. Chen, W. Cheng, and C. Wang. Multi-aperture overlap-scanning technique for large aperture test. In *Laser Interferometry IV: Computer-Aided Interferometry*, volume 1553, pages 626–635. SPIE, SPIE, 1991.

[65] J. Fleig, P. Dumas, P. E. Murphy, and G. W. Forbes. An automated subaperture stitching interferometer workstation for spherical and aspherical surfaces. In *Advanced characterization techniques for optics, semiconductors and nanotechnologies*, number 5188 in Proceedings of SPIE, 2003.

[66] M. Bray. Stitching interferometer for large plano optics using a standard interferometer. In *Optical Manufacturing and Testing II*, volume 3134 of *Proceedings of SPIE*, 1997.

[67] M. Bray. Stitching interferometry: side effects and PSD. In *Optical Manufacturing and Testing III*, volume 3782 of *Proceedings of SPIE*, 1999.

[68] M. Bray. Interférométrie à couplage de zones: Progrès récents. In *Méthodes et Techniques Optiques pour l'Industrie*, Deuxième colloque francophone: SFO / CMOI, 2003.

[69] M. Bray. Stitching interferometry and absolute surface shape metrology. In *Optical Manufacturing and Testing IV*, volume 4451 of *Proceedings of SPIE*, 2001.

[70] P. J. Sullivan, R. E. Parks, and L. Shao. Stitching of equatorial profiles for extended spatial range assesment. In *Specification, Production and Testing of Optical Components and Systems*, volume 2775, pages 67–73. SPIE, May 1996.

[71] P. Murphy, G. Forbes, J. Fleig, P. Dumas, and M. Tricard. Stitching interferometry: a flexible solution for surface metrology. *Optics and Photonics News*, pages 38–43, May 2003.

[72] M. Tricard and P. Murphy. Subaperture stitching for large aspheric surfaces. NASA Space Optics Manufacturing Technology Centre: Tech Days: contributed presentation. http://optics.nasa.gov/tech_days/tech_days_2005/, August 2005. Retrieved 22/01/07.

[73] Personal communication with UPS group, Optic Technium N. Wales., 2007.

[74] Personal communication with Dr. Andrew Lewis of NPL on refractive index changes in columns of air., 2007.

[75] D. D. Walker. Personal communication, 2007.

[76] OMAM. OMAM project meeting decision. Minutes., 2002.

[77] The MathWorks. *Optimization Toolbox for use with Matlab.* The MathWorks, 5th ed. for ver. 3.0.2 edition, 2005.

[78] G. Peggs. Method for measuring objects. Patent application, May 2004.

[79] M. J. Callender. *Hybridization of the swing arm profilometer with a wavefront sensor.* PhD thesis, University College London, 2007. In progress.

[80] Claude Roddier and François Roddier. Wave-front reconstruction from defocused images and the testing of ground-based optical telescopes. *J. Opt. Soc. Am. A*, 10(11):2277–2287, November 1993.

[81] M. Callender and C. King. Arden AWS-50 wavefront sensor evaluation. Technical report, Optical Science Laboratory, University College London, 2004. Commercial in confidence.

[82] A. H. Greenaway. Personal communication with A. H. Greenaway, 2003.

[83] X. Levecq. Wavefront sensors. http://www.dmoptics.com/wavefront.htm, March 2007.

[84] D. D. Walker. Personal communication, 2005.

[85] David Walker. Personal communication, 2004.

[86] T. Yano, N. Gouda, Y. Kobayashi, T. Tsujimoto, T. Nakajima, H. Hanada, and Y. Kan-ya. Ccd centroiding experiment for the Japan astrometry satellite mission (JASMINE) and in-situ lunar orientation measurement. *Astronomical Society of the Pacific*, 116:667–673, July 2004.

[87] W. Cheng and M. Chen. Transformation and connection of subapertures in the multiaperture overlap-scanning technique for large optics tests. *Optical Engineering*, 32(8):1947–1950, August 1993.

[88] M. Sjödahl and B. F. Oreb. Stitching interferometric measurement data for inspection of large optical components. *Optical Engineering*, 41(2):403–408, February 2002.

[89] J. Negro. Subaperture optical system testing. *Applied Optics*, 23(12):1921–1930, June 1984.

[90] E. W. Weisstein. From MathWorld – Genus. http://mathworld.wolfram.com/Genus.html, August 2006.

[91] E. W. Weisstein. From MathWorld – Affine Transformation. http://mathworld.wolfram.com/AffineTransformation.html, August 2006.

[92] J. Foley, A. van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice.* Addison-Wesley, 2nd edition, 1990.

[93] E. W. Weisstein. From MathWorld – Rodrigues rotation. http://mathworld.wolfram.com/RodriguesRotationFormula.html, August 2006.

[94] L. D. Landau and E. M. Lifshitz. *Mechanics.* Butterworth Heinemann, 1976. Trnaslated by Sykes and Bell.

[95] E. W. Weisstein. From MathWorld – Euler Angles. http://mathworld.wolfram.com/EulerAngles.html, August 2006.

[96] L. Allen *et al.* Technical report synopsis: Hubble space telescope optical systems failure report. Technical report, Lockheed Martin Corp., November 2006.

[97] S. Chen, S. Li, Y. Dai, and Z. Zhen. Lattice design for subaperture stitching test of a concave paraboloid surface. *Applied Optics*, 45(10):2280–2286, April 2006.

[98] R. Fletcher. *Practical Methods of Optimization.* Wiley Interscience. John Wiley and Sons, 2nd edition, 1987.

[99] W. Murray, editor. *Numerical Methods for Unconstrained Optimization.* Academic Press, 1972.

[100] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in Fortran,* pp 51–63. Cambridge, 2nd edition, 1992.

[101] The MathWorks. *Optimization Toolbox for use with Matlab.* The MathWorks, 5th ed. for ver. 3.0.2 edition, 2005.

[102] Wikipedia. Procrustes. http://en.wikipedia.org/wiki/Procrustes, February 2007.

[103] G. A. F. Seber. *Multivariate Observations.* Wiley, 1984.

[104] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in Fortran.* Cambridge, 2nd edition, 1992.

[105] Personal communication with Dr. Richard Leach of NPL concerning filtering and form removal., 2007.

[106] 4d Technology corp's 4Sight software: www.4dtechnologycorp.com, 2007.

[107] Personal communication with Dr. Richard Leach on 3D surface texture and form standards and the filtering of data., 2007.

[108] L. A. Selberg. Interferomter accuracy and precision. In *Proceedings of SPIE: Optical Fabrication and Testing,* volume 1400, pages 24–32. SPIE, 1990.

[109] Personal communication with Zeeko Ltd, the project's industrial sponsors, coalville, leics, 2007.

[110] D. Anderson, R. Parks, and L. Shao. Versatile profilometer for aspheric optics. *Technical Digest on Optical Fabrication and Testing,* 11:119–122, 1990.

[111] David S. Anderson and James H. Burge. Swing-arm profilometry of aspherics. In Victor J. Doherty and H. Philip Stahl, editors, *Optical Manufacturing and Testing*, volume 2536, pages 169–179. SPIE, 1995.

[112] M. J. Callender, A. Efstathiou, C. W. King, D. D. Walker, A. E. Gee, A. J. Lewis, S. Oldfield, and R. M. Steel. Swing arm profilometer for large telescope mirror element metrology. In Eli Atad-Ettedgui, Joseph Antebi, and Dietrich Lemke, editors, *Optomechanical Technologies for Astronomy*, volume 6273, page 62732R. SPIE, 2006.

[113] C. King, M. Callender, A. Efstathiou, D. Walker, A. Gee, A. Lewis, S. Oldfield, and R. Steel. A 1 m swing-arm profilometer for large telescope mirror element metrology. In *Proceedings of the 2006 ASPE Annual general Meeting.* American Society for Precision Engineering, October 2006. Paper 1913.

[114] A. Lewis, S. Oldfield, M. Callender, A. Efstathiou, C. King, and D. Walker. Arcuate arm profilometry - traceable metrology for large mirrors. In *Simposio de Metrologia 2006.* CENAM, October 2006.

[115] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in Fortran*, pp 18–21. Cambridge, 2nd edition, 1992.

[116] Zygo Corporation. Transmission sphere selection. Technical report, Zygo corporation, www.zygo.com, 2005.

[117] David Brooks. Personal communication, 2006. Optical Science Laboratory, UCL.

[118] R. E. Parks and C. Evans & L. Shao. Calibrating interferometer transmission spheres. In *Optical Society of America, Optical Fabrication and Testing Workshop, Hawaii*, 1999.

[119] Zygo Corporation. Product borchure for GPI interferometers., 2007.

[120] Wyko Vision software manual, Wyko Corp, 2005.

[121] R. E. Parks. A practical implementation of the random ball test. In *Optical Soceity of America Optical Fabrication and Testing Meeting.* OSA, 2006.

[122] U. Griesmann, Q. Wang, J. Soons, and R. Carakos. A simple ball averager for reference sphere calibrations. Technical report, NIST, 2006.

APPENDIX A

Glossary of Terms

A.1 Acronyms and Abbreviations

CCD	Charge Coupled Device
CGH	Computer Generated hologram
CMM	Coordinarte Measuring Machine
ELT	Extremely Large Telescope
λ	Measurement wavelength (usually 632.8 nm)
LVDT	Linear Variable Differential Transformer
NPL	National Physical Laboratory
NURBS	Non-Uniform Rational B-Spline
OPD	Optical Path Difference
P-V	Peak-to-Valley
RMS	Root Mean Square
RoC	Radius of Curvature
SAP	Swing-Arm Profilometer
SUT	Surface Under Test
WFS	WaveFront Sensor
PDWFS	Phase Diversity WaveFront Sensor

A.2 Mathematical Symbols

$\mathbf{a}$	A vector
$\mathbf{B}$	A matrix
$\mathbb{R}^n$	Set of reals in n-dimensional space
$A,\ B,\ \psi,\ x,\ y,\ z$	Data transformation parameters
$\mathbf{t}_i$	Vector of transformation parameters for the i^{th} sub-aperture
$\mathbf{T}_i$	6-parameter transformation matrix for the i^{th} sub-aperture
$\{\ldots\}$	A set
$\mathrm{Det}(\mathbf{A})$	Determinant of $\mathbf{A}$
$\mathbf{A}^{-1}$	Matrix inverse of $\mathbf{A}$
$\mathbf{A}^T$	Matrix or array transpose of $\mathbf{A}$
$\bar{r}$	mean value of $r_1 \ldots r_n$
$\|\mathbf{x}\|$	norm of $\mathbf{x}$
$\overset{*}{\mathbf{x}}$	Optimal solution to minimization problem
$\mathbf{x}^{(k)}$	Values for $\mathbf{x}$ at the k^{th} iteration
$\neq$	Not equal to
$\forall$	For all
$\in$	An element of
${}_nC_r$	The binomial coefficient, n choose r.
$[\mathbf{U}\ \mathbf{S}\ \mathbf{V}] = \mathrm{SVD}(\mathbf{A})$	The singular value decomposition of $\mathbf{A}$
ΔA	Uncertainty or error in A

A.3 List of Indices

i Sub-aperture number

n Total number of sub-apertures

Appendix B

Publications

C. W. King, M. J. Callender, A. Efstathiou, D. D. Walker, A. E. Gee, A. J. Lewis and S. Oldfield. A 1 m Swing-Arm Profilometer for Large Telescope Mirror Elements. *In Proceedings of 2006 ASPE AGM*. American Society for Precision Engineering, Oct. 2006. Paper 1913.

A. Lewis, S. Oldfield, M. J. Callender, A. Efstathiou, A. E. Gee, C. W. King, D. D. Walker. Arcuate Arm Profilometry - Traceable Metrology for Large Mirrors. *In Proceedings of Mexican Simposio de Metrologia, 2006*. CENAM, Oct. 2006.

M. J. Callender, A. Efstathiou, C. W. King, D. D. Walker, A. E. Gee, A. J. Lewis and S. Oldfield. Swing-Arm Profilometer for Large Telescope Mirror Element Metrology. *Proceedings of SPIE: Optomechanical Technologies for Astronomy.* Volume 6273, p6273R, SPIE 2006.

APPENDIX C

UCL Prototype PDWFS

C.1 UCL Imaging System

Two weeks spent by the author and a colleague at Heriot-Watt resulted in the construction of a prototype for the mechanical components of a PDWFS. The baseline specification for the optical components and mechanical interface could be derived from this model and used in the design of the swing-arm profilometer. A baseline specification, drawn up with input from Heriot-Watt is shown in table C.1

A simple imaging system comprising two achromatic doublets was concieved using thin-lens theory. Construction was carried out using commerical off-the-shelf components. Achromatic lenses, chosen for their higher quality than equivalent singlets, were mounted in fine focussers to allow for the approximations made in the thin-lens equations and the overall magnification of $\times 1/60$ resulted in an order size of approximately 40 pixels on the CCD camera. Figure C.1 shows an image produced by the system. The zeroth order, at centre, is an image of the point source that provides the test wavefront. The two larger circular features are the defocussed images of the point source in the ± 1 orders of the grating used. The optical

Aperture Size	15 mm ⌀
Nominal object lens – SUT distance	350 mm
Overall magnification	$\times\frac{1}{60}$
Nominal order size on CCD	250 μm ⌀
CCD pixel size	8.3 × 8.3 μm
CCD	782 × 582 pixels, 8 bits
Operating wavelength	632.8 nm (He-Ne)
Object lens	18 mm ⌀, f = 50 mm achromat
Re-imaging lens	10 mm ⌀, f=10 mm achromat

Table C.1: Basic specifications for the prototype OMAM (H-W/UCL) wavefront sensor hardware.

Figure C.1: Defocus images of a point source produced by the device shown in figure C.3

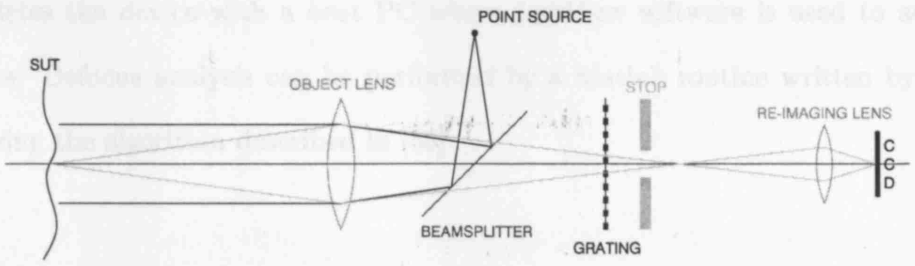


Figure C.2: Optical schematic of the H-W/UCL prototype wavefront sensor.

schematic is shown in figure C.2 and the design allows the grating to be changed readily

Figure C.3: WFS: Form and main components.

The imaging system plane-wave illumination source is derived from a single-mode (TEM_{00}) 10 μm $\varnothing$ optical fibre operating at 633 nm. A microscope objective lens is used to launch the light from a 1 mW laser down the fibre. The point source, which is located at a object lens conjugate, is introduced into the system via a cube beamsplitter to give a collimated beam. A translation stage allows the point source's position in a plane perpendicular to the optical axis to be adjusted, and a kinematic mount serves to adjust the beamsplitter. A firewire camera interface completes the device with a host PC where LabView software is used to acquire images. Defocus analysis can be performed by a Matlab routine written by H-W involving the algorithm described in [55]

Figure C.4: WFS: Elevation and section showing key optical components.

APPENDIX D

Prototype WFS Model Drawings

APPENDIX E

Fisba File Format

This short C program was written to convert a Fisba interferometer phasemap into a format that can be read by Matlab. The Fisba binary file format is as follows:

Type (byte 1): a 16 bit signed integer specifying the data as phase in Angstroms

Rows (bytes 2–5): a 32 bit signed integer specifying the number of rows in the phase data array.

Columns (bytes 6-9): a 32 bit integer specifying the number of columns in the phase data array.

Not Aperture (bytes 10-13): a 32 bit singed integer specifying the data value to be regarded as not aperture data.

Comment length (bytes 14–17): a 32 bit singed integer specifying the length of the user supplied comment for the data file.

Comment (from byte 18 to 18+comment length): The user supplied comment of length specified by *Comment Length* above.

Phase Data (remaining bytes): An array of the phase data of size rows×columns where each value is a 32 bit signed integer representing the phase value in Angstroms.

The software reads in a Fisba format file and outputs the type, rows, columns, not aperture, comment length and phase data into ASCII format that can be piped to a file for subsequent input into Matlab. The arithmetic shifts (<<n) move the bits in a double word by n places to the left. n=1 is equivalent to multiplying the value by 2. The shifts are used to compose 32-bit signed integers in a 32 bit CPU register from the 4 8-bit singed integers read for each value in question.

```
main()
{
/* Create the file pointers and declare the variables */
        FILE *fp;
        char path[255];
        int type, n;
        long  x, y, rows, cols, not_aperture, comment, a, b, c, d, data;
/* Prompt user for the filemane to open and open the file
   if it exists */
        scanf("%s", &path);
        if( (fp=fopen( path,"rb"))!=NULL)
        {
/* Read from file pointer into type variable and output to stdout */
                type=fgetc(fp);
                printf("#Type = %d\n", type);
/* Read 4 8-bit integers and convert them to the 32 bit form by
   arithmetic shifts appropriate to each position in binary form
   for the rows, columns, not-aperture and comment length.
   Output to stdout */
                a=(long)fgetc(fp);
                b=(long)fgetc(fp)<<8;
                c=(long)fgetc(fp)<<16;
                d=(long)fgetc(fp)<<24;
                rows=a+b+c+d;
                printf("#Rows = %ld\n", rows);
                a=(long)fgetc(fp);
                b=(long)fgetc(fp)<<8;
                c=(long)fgetc(fp)<<16;
                d=(long)fgetc(fp)<<24;
```

```
                cols=a+b+c+d;
                printf("#Cols = %ld\n", cols);
                a=(long)fgetc(fp);
                b=(long)fgetc(fp)<<8;
                c=(long)fgetc(fp)<<16;
                d=(long)fgetc(fp)<<24;
                not_aperture=a+b+c+d;
                printf("#Not_aperture = %ld\n", not_aperture);
                a=(long)fgetc(fp);
                b=(long)fgetc(fp)<<8;
                c=(long)fgetc(fp)<<16;
                d=(long)fgetc(fp)<<24;
                comment=a+b+c+d;
                printf("#Comment length = %ld\n", comment);
/* Read the comment and do nothing with it */
                for(n=0;n<comment;++n)
                {
                        fgetc(fp);
                }
/* Read the phase data 4 8-bit integers at a time and compose
   the 32 bit value by means of arithmetic shifts again. Do
   two's complement arithmetic to compose correct neg values */
                for(x=0;x<rows;++x)
                {
                        for(y=0;y<cols;++y)
                        {
                                a=(long)fgetc(fp);
                                b=(long)fgetc(fp)<<8;
                                c=(long)fgetc(fp)<<16;
                                d=(long)fgetc(fp)<<24;
                                if(d>2147483647)
                                {
                                        data=~(a+b+c+d) +1;
                                }
                                else data = a+b+c+d;
                                printf("%ld ", data);
                        }
                        printf("\n");
                }
```

```
/* Close the file */
                fclose(fp);
        }
        else
/* Prints file not found if user makes a mistake */
                printf("File not found\n");
}
```

E.1 Calculation of the Angle ϕ_{sweep}

The problem of how to calculate the angle ϕ_{sweep} can be simplified greatly if one imagines looking at the swing-arm profilometer side-on when the swing arm is positioned such that ϕ is equal to zero and the tilt angle of the swing-arm is equal to θ. See figure E.1. That is one looks at an elevation of the swing-arm when correctly set up with the effective arm length equal to l and the radius of curvature of the part under test equal to R. Firstly, it is important to realize that the circumference of the part under test will project a single line segment onto this elevation with a length equal to the diameter of the part under test, D. In Cartesian co-ordinates with the origin at the centre of curvature of the part under test, this line segment is part of the general line with the equation $y = const$. Now we imagine the circle traced out

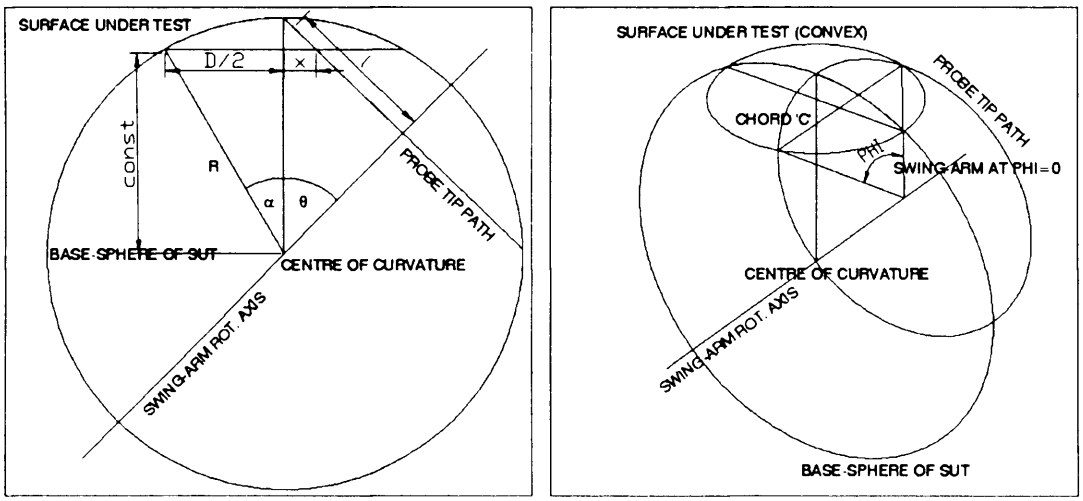


Figure E.1: Side (elevation) and oblique views of the arrangement.

by the tip of the profilometer as the arm is swung around through one complete revolution. Again, this circle when viewed from a side elevation is seen as a second line segment with its length equal to two times the effective arm length, l. If the profilometer has been set up correctly then this line will intersect with the previous line at the points where the profilometer tip falls off the side of the part under test. The equation of this line can be easily realized. The second line passes through the point $(0, R)$, where R is the radius of curvature of the part under test, and has a gradient equal to $\tan\theta$[1]. Therefore, the equation of the line is $y = (\tan\theta)\,x + R$, since when $x = 0$, $y = R$. One has to merely find the x-co-ordinate of the intersection of these two lines to help deduce the single chord length shown for each of the two circles in the diagram. When the chord length C is known, the value of ϕ can be calculated directly. The intersection occurs when $y = const$ for both lines. Referring to figure E.1 on page 319, and using basic trigonometry we see that ,

$$const = R\cos\left(\sin^{-1}\left(\frac{D}{2R}\right)\right). \tag{E.1}$$

We have the equation of our second line,

$$y = x\tan\theta + R. \tag{E.2}$$

When the lines intersect, we are at the point $(x, const)$, therefore,

$$R\cos\left(\sin^{-1}\left(\frac{D}{2R}\right)\right) = x\tan\theta + R. \tag{E.3}$$

Solving for x yields,

$$x = \frac{R\cos\left(\sin^{-1}\left(\frac{D}{2R}\right)\right) - R}{\tan\theta} \tag{E.4}$$

Now, using a plan view of the profilometer, figure E.2 on page 321, we can see that,

$$\left(\frac{D}{2}\right)^2 = x^2 + \left(\frac{C}{2}\right)^2, \tag{E.5}$$

[1]There is a minus sign here but it can be ignored due to the symmetry of the geometry and the fact that we will be squaring an expression involving $\tan\theta$

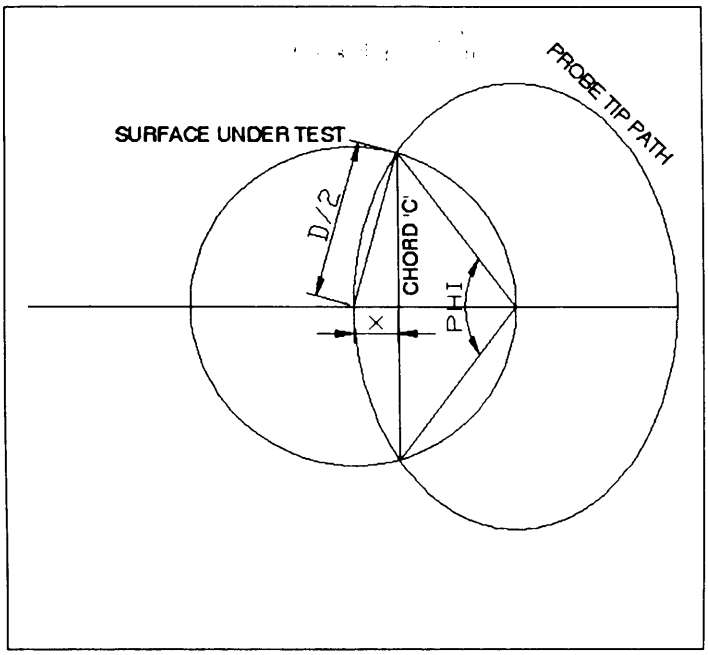


Figure E.2: Plan view of the swing-arm and SUT arrangement.

where C is the chord length required. Therefore,

$$C = 2 \cdot \sqrt{\left(\frac{D}{2}\right)^2 - x^2}. \tag{E.6}$$

Substituting all the numbers in yields,

$$C = 2 \cdot \sqrt{\left(\frac{D}{2}\right)^2 - \left(\frac{R\cos\left(\sin^{-1}\left(\frac{D}{2R}\right)\right) - R}{\tan\theta}\right)^2} \tag{E.7}$$

The value of θ is calculated in the usual manner for the swing-arm profilometer i.e. $\sin\theta = \frac{l}{R}$. To calculate the value of ϕ_{sweep}, we see that the arm-length is equal to l as previously stated and that the ends of the swing-arc for the arm are the ends of the chord just calculated. We therefore have an isosceles triangle as shown in figure E.3 on page 322. Therefore we arrive at the value for ϕ_{sweep},

$$\phi_{sweep} = 2 \cdot \sin^{-1}\left(\frac{C}{2l}\right). \tag{E.8}$$

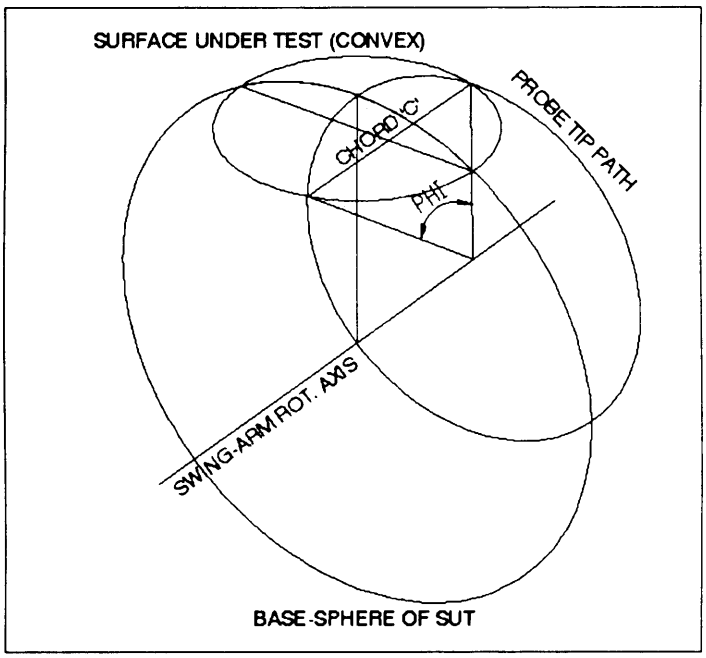


Figure E.3: View of the triangle used to calculate ϕ.

APPENDIX F

The Annulus Centroid Method

An annulus is placed on the phasemap at the centroid of each fiducial in turn and the phase values lying within the annulus are averaged to elucidate the z-coordinate of the fiducial centroid. The program is implemented in Matlab as a sub-window of the main stitching software GUI. The sequence of events is as follows:

1. The phasemap already loaded into memory is displayed in a plot window.
2. The software prompts the user to pick a point on the phasemap inside the first fiducial.
3. A search is initiated inside a circular region larger than the fiducial for data points containing the NaN constant. The average value of the x- and y-coordinates of the NaN conatants is found. These become the x- and y-coordinates of the fiducial centroid.
4. The phase values under an annulus centred on the fiducial centroid location are averaged to yield the z-coordinate of the fiducial.
5. The process is repeated as many times as necessary fom point (2) above.

6. The coordinates of all the fiducials are appended to the sub-aperture array data in the order that they were picked

```
% ********* Fid acquire **********
% Prompt user to pick  points inside and outside a fiducial
a1=round(ginput(1));
b=round(ginput(1));
% compute the radius of the circle centred
% at point a1 and passing through b.
rad=sqrt((b(1)-a1(1)).^2 + (b(2)-a1(2)).^2);
% Plot the search circle on the phaseplot
hold on
plot(a1(1),a1(2),'ro');
n=1;
q=0:pi/60:2*pi;
circ=[rad*cos(q)+a1(1,1) rad*sin(q)+a1(1,2)];
circ=reshape(circ,size(q,2),2);
plot(circ(:,1), circ(:,2),'r');
hold off
% Set up variables for computing the (x, y) centroid
num=0;
x=0;
y=0;
hold on
% Search inside the search circle for NaN points and average the coords
% Also colour the points yellow on the phasemap
for i=1:size(a.WykoValues,1)
    for j=1:size(a.WykoValues,2)
        if  (i-a1(2)).^2 + (j-a1(1)).^2 < rad.^2 & isnan(a.WykoValues(i,j))
            x=x+j;
            y=y+i;
            plot(j,i,'y');
            num=num+1;

        end
    end
end
hold off
set(handles.x,'String', num2str(x/num));
```

```
set(handles.y,'String', num2str(y/num));
set(handles.num,'String',num2str(num));
% The xf and yf are the x- and y-coordinates of the fiducial
xf(1)=x/num;
yf(1)=y/num;
% Now determine z by the annulus method
mark_1=0;
num=0;
% Search inside the annulus centred around xf(1),yf(1) and average the phase
for i=1:1:size(a.WykoValues,1)
    for j=1:1:size(a.WykoValues,2)
        if (i-yf(1)).^2 + (j-xf(1)).^2 > 25 & (i-yf(1)).^2 + (j-xf(1)).^2 < 49
            mark_1=mark_1 + a.WykoValues(i,j);
            a.WykoValues(i,j)=NaN;
            num=num+1;
        end
    end
end
% mark_1 is the z-coordinate of the fiducial
mark_1=mark_1/num;
% ********** End of fid acquisition *********
```

APPENDIX G

The Optimizer Object Function

The object function is at the core of the optimization process and therefore the code to compute f is revealed here in all its splendour. The object function accepts several variables and returns a scalar value.

Inputs

PA: The initial guess values for each of the transformation matrix parameters formatted as an array of double precision floats.

$$\mathbf{PA} = \begin{pmatrix} A_1 & B_1 & \psi_1 & x_1 & y_1 & z_1 \\ A_2 & B_2 & \psi_2 & x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_n & B_n & \psi_n & x_n & y_n & z_n \end{pmatrix}, \tag{G.1}$$

where n is the number of sub-apertures to stitch. The optimizer determines how many sub-apertures by interrogation of the number of rows of this matrix.

fid_matrix: This is the fiducial matrix as defined in figure 3.9 on page 155. For

example, consider the 2×2 matrix,

$$\begin{pmatrix} 4 & 32 \\ 21 & 0 \end{pmatrix}. \tag{G.2}$$

This indicates to the object function that fiducial number 1 (row 1 in the matrix G.2) occurs in sub-apertures 1 and 2 and it's coordinates can be found in column numbers 4 and 32 respectively. Fiducial number 2 is only located in sub-aperture 1 and its coordinates can be found in column number 21 of the sub-aperture data. The zero indicates that fiducial number 2 does not occur in sub-aperture 2 (i.e. indicated by row 2 of this matrix). Therefore for 12 fiducials in 3 sub-apertures we shall have a 12×3 matrix.

sub_array: This is a Matlab cell array containing the sub-aperture data, one set for each cell. A sub-aperture array is called `sub_array{i}`, where i is the i^{th} sub-aperture and `sub_array{1}(:,21)` returns a 4×1 column vector $(x, y, z, 1)^T$ which referring to G.2 above corresponds to the coordinates of fiducial number 2 in sub-aperture number 1.

Output

The object function outputs a scalar which is the value of the object function. **The code is totally independent of the number of sub-apertures to stitch** and can therefore be used to stitch any number of sub-apertures that can be stored in the computer's memory space.

Code

```
% object_3d_1.m
% A constrained minimization based on the parameters fed into the
% optimizer. See code for some lovely loops. This function computes
% the sum of all the distances squared between all pairs of fiducials.
```

```
% The optimizer then attempts to minumize this sum.
function f = object_3d_1(PA, fid_matrix, sub_array)
% Loopt and compute the rotation matrices T_i.
for n=1:size(sub_array,2)
    if size(sub_array{n},2) == 0
% Decision used to skip absent sub-apertures
        transform{n}=[0 0 0 0; 0 0 0 0 ; 0 0 0 0 ; 0 0 0 0 ];
    else
% The transformation matrix
transform{n}=...
        [(cos(PA(n,3))) (-sin(PA(n,3))) 0
(PA(n,4).*cos(PA(n,3)) - PA(n,5).*sin(PA(n,3)));
        (sin(PA(n,3))) (cos(PA(n,3))) 0
(PA(n,4).*sin(PA(n,3)) + PA(n,5).*cos(PA(n,3)))
        PA(n,1) PA(n,2) 1 0;
        0 0 0 1];
    end
end
% Transform the points using a very inefficient nested loop
for fiducial_index=1:size(fid_matrix,1)
    for sub_index=1:size(fid_matrix,2)
        if fid_matrix(fiducial_index,sub_index) == 0
        else
            sub_array{sub_index}(1:4,fid_matrix(fiducial_index,sub_index)) = ...
             transform{sub_index}*sub_array{sub_index}(1:4,fid_matrix(fiducial_index,sub_index));
        end
    end
end
% Now calculate the distance between the fiducials.
% Initialize the cumulative totals for the x, y, and z sums
partial_sum_x = 0;
partial_sum_y = 0;
partial_sum_z = 0;

% Now parse the fiducial matrix and determine the fiducial pairings at
% the same time as computing the 3-D distance between pairs.

for fiducial_index=1:size(fid_matrix,1)
    for sub_index=1:size(fid_matrix,2)
```

```
        if fid_matrix(fiducial_index,sub_index) == 0
        else
            for inc_index=sub_index+1:size(fid_matrix,2)
                if fid_matrix(fiducial_index,inc_index) == 0
                else
                    % Now we sum
                    partial_sum_x = partial_sum_x +
                    (sub_array{sub_index}(1,fid_matrix(fiducial_index,sub_index)) -
                    sub_array{inc_index}(1,fid_matrix(fiducial_index,inc_index))).^2;

                    partial_sum_y = partial_sum_y +
                    (sub_array{sub_index}(2,fid_matrix(fiducial_index,sub_index)) -
                    sub_array{inc_index}(2,fid_matrix(fiducial_index,inc_index))).^2;

                    partial_sum_z = partial_sum_z +
                    (sub_array{sub_index}(3,fid_matrix(fiducial_index,sub_index)) -
                    sub_array{inc_index}(3,fid_matrix(fiducial_index,inc_index))).^2;
                end
            end
            break % Next please!
        end
    end
end
% Form the sum. The partial sums contain the squared distances in x, y and z between fiducial pairs.
f = partial_sum_x + partial_sum_y + partial_sum_z;
```

Phew!

APPENDIX H

Calling `fmincon.m` and `fminunc.m`

The Matlab optimization routine is called by the stitching software and supplied with a set of options to control how it runs. The calling convention for `fminunc.m` is:

```
[t, fval, exitflag, output] = ...
fmincon('object_3d_1', PA, [], [], [], [], lb, ub, [],...
   options, fid_matrix, sub_array);
```

Inputs

'object_3d_1.m': The name of the object function to minimize is passed to the optimizer such that it can be called when necessary at each iteration. `object_3d_1.m` is described in appendix G.

PA: This is the array of initial values for the transformation matrices. Its format is descibed in appendix G.

lb: The lower bounds for the transformation parameters used in a constrained optimization: $[A_l, B_l, \psi_l, x_l, y_l, z_l]^T$.

ub: The upper bounds for the transformation parameters used in a constrained optimization: $[A_u, B_u, \psi_u, x_u, y_u, z_u]^T$.

options: A structure containing the option values to be passed to the optimizer.

```
options = optimset('Display', 'iter',
'MaxFunEvals', n, 'MaxIter',m, 'TolFun', tf,
'TolX', tx, 'Diagnostics', 'on', 'DiffMa', dmx, 'DiffMi',dmn);
```

These options will cause the optimizer to display progress at each iteration, make a maximum of n function evaluations or m iterations. The tolerance on the object function is set to tf, the tolerance on x to tx, the diagnostic information is turned on and the maximum and minimum changes in the perturbation for the finite differencing are set to dmx and dmn respectively.

fid_matrix: This is the fiducial matrix as defined in appendix G and elsewhere.

sub_array: This is a Matlab cell array containing the sub-aperture data. see appendix G and the main text for more details.

Outputs

t: The set of optimal transfomation parameters for the stitching in the form:

$$\mathbf{t} = \begin{pmatrix} A_1 & B_1 & \psi_1 & x_1 & y_1 & z_1 \\ A_2 & B_2 & \psi_2 & x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_n & B_n & \psi_n & x_n & y_n & z_n \end{pmatrix}, \tag{H.1}$$

where n is the number of sub-apertures to stitch.

fval: The object function value after the optimizer terminates.

exitflag: The reason why the optimizer terminated. Value as follows:

1 Magnitude of gradient smaller than the specified tolerance.

2 Change in x was smaller than the specified tolerance.

3 Change in the object function value was smaller than the specified tolerance.

0 Number of iterations exceeded `options.MaxIter` or number of function evaluations exceeded `options.MaxFunEvals`.

-1 Algorithm was terminated by the output function. (Output function not used in this work)

-2 Line search cannot find an acceptable point along the current search direction.

output: A structure containing various statistics concerning the optimization.

output.iterations: The number of iterations reached when the optimizer terminated.

output.funcCount: The number of object function evaluations reached when the optimier terminated.

output.stepSize: The stepsize at the last iteration.

firstOrderOptimality: a measure of the optimality of the process in terms of the infinity norm of the gradient of the object function at the last point.

algorithm: The algorithm used by the optimizer. In this work, this is always a medium-scale quasi-Newton line search. See section 4.5 and [101] for details.

message: A message such as 'Optimizer terminated successfully: relative infinity norm of gradient less than options.grad'. The message selected depends upon the reason the optimizer terminated.

iter_data: This is not standard in the Matlab optimization toolbox function, `fminunc.m`. The code has been modified to produce a step by step history

of the optimization. The output is a vector containing the function count, the object function value, the step size, and the first order optimality referenced by iteration number such that analysis plots can be made as shown in figure 6.5 on page 232.

APPENDIX I

Interferogram Reconstruction

This code works by interfering two wavefronts and computing the resultant amplitude. One wave is the reference wave with a constant phase, the other is the measured wavefront. The amplitused of each are computed over the phase-array and then the precomputed tip/tilt is added.

```
% Enter the tip, tilt and piston required to be added
% to the interferogram here
tip=0;
tilt=0;
piston=0;
x=1;
% set the wavelength
lambda=633;
% Read the values from the phasemap from stitching software.
load Y
% Define an empty array the same size as the phasemap just read in to
% memory
t=zeros(size(Y,1),size(Y,2));
% Now go to each pixel in 't' in turn and apply the tip/tilt to the data by adding
% a constant (tip/tilt) multiplied by x and y coordinates scaled by the
% aperture
for i=1:size(Y,1)
    for j=1:size(Y,2)
        t(i,j)=Y(i,j) + tip*(j*633)/size(Y,2) + tilt*(i*633)/size(Y,1);
```

```
    end
end
% Now form the fringes by computing the interference as a function of
% the air gap (phase received at detector * wedge) and add the piston etc.
sum=sin(2*pi*x/lambda)+sin(2*pi*(((2*(t+(316.5*piston))) + (lambda/2) + x))/lambda);
% Output the data.
figure
surf(xi,yi,sum)
view(0,90)
axis equal
colormap gray;
```

APPENDIX J

Automatic Centroiding Code

J.1 Matlab Calling Code

```
% First load the LabView compiled DLL
loadlibrary omam_proto_lib omam_proto_lib;
% Declare a variable called data to store the image data from Labview
data=uint8(zeros(55460,1));
% Decalre a variable to store the centroids from Labview
cen=double(zeros(60,1));
%Declare a pointer to pass to the Labview DLL
% to indicate where to store the image data
pdata=libpointer('uint8Ptr',data);
% Declare a pointer to pass to Labview DLL
% to indicate where to store the centroid data
pcen=libpointer('doublePtr',cen);
% Call the LabView DLL passing the pointers and array sizes
[data cen]=calllib('omam_proto_lib','Centroid',pdata,55460,pcen,60);
%Re-shapre the arrays into  square ones
data=reshape(data,235,236)';
cen=reshape(cen,2,30)';
% Unload the LabView DLL
unloadlibrary omam_proto_lib;
% Plot the image data and the centroids.
```

```
imagesc(data);
axis image;
hold on;
for n=1:size(cen,1)
    plot(cen(n,1),cen(n,2),'b+');
end
hold off;
colormap gray
```

J.2 LabView VI